

Publication No. US20030049619A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Maris, Barry
TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Libraries Of Polynucleotides
FILE REFERENCE: HER0041
CURRENT APPLICATION NUMBER: US/09/813,408
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 440
TYPE: PRT
ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match 64.8%; Score 35; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
|:|:| |:|
Db 120 EVLPWGVDY 128

RESULT 7
US-09-815-242-5111
Sequence 5111, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5111
LENGTH: 1062
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5111

Query Match 63.0%; Score 34; DB 10; Length 1062;
Best Local Similarity 85.7%; Pred. No. 2,4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
|:|:| |:|
Db 321 PQGMDYS 327

Publication No. US20030049619A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Maris, Barry
TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Libraries Of Polynucleotides
FILE REFERENCE: HER0041
CURRENT APPLICATION NUMBER: US/09/813,408
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 440
TYPE: PRT
ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match 64.8%; Score 35; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
|:|:| |:|
Db 120 EVLPWGVDY 128

RESULT 7
US-09-815-242-5111
Sequence 5111, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5111
LENGTH: 1062
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5111

Query Match 63.0%; Score 34; DB 10; Length 1062;
Best Local Similarity 85.7%; Pred. No. 2,4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
|:|:| |:|
Db 321 PQGMDYS 327

RESULT 8
US-10-027-806-4
Sequence 4, Application US/10027806
Patent No. US20020160476A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCCP.002A
CURRENT APPLICATION NUMBER: US/10/027,806
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3472
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
|:|:| |:|
Db 2294 EDVIPRGISFS 2304

RESULT 9
US-10-034-623-4
Sequence 4, Application US/10034623
Publication No. US20020198365A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCCP.002A
CURRENT APPLICATION NUMBER: US/10/034,623
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/102,294
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3472
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
|:|:| |:|
Db 2294 EDVIPRGISFS 2304

RESULT 10
US-10-027-801-4
Sequence 4, Application US/10027801
Publication No. US20030054364A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.

```

; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCRP.002A
; CURRENT APPLICATION NUMBER: US/10/027,801
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

```

```

Query Match      63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMDYS 11
Db 2294 EDVIPRGISFS 2304

```

RESULT 11

```

US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102702A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 2001-09-06
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

```

```

Query Match      61.1%; Score 33; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 15;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMDYS 11
Db 38 EXHIFGLEY 48

```

RESULT 12

```

US-09-738-626-4881
; Sequence 4881, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

```

```

; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4881
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4881

```

```

Query Match      61.1%; Score 33; DB 9; Length 283;
Best Local Similarity 75.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 VPXGMDYS 11
Db 56 VPAGADYS 63

```

RESULT 13

```

US-09-815-242-10697
; Sequence 10697, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10697
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10697

```

```

Query Match      61.1%; Score 33; DB 10; Length 299;
Best Local Similarity 40.0%; Pred. No. 94;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMDY 10
Db 218 EQITPTGIEY 227

```

us-09-909-164-13.rapb

Wed Jun 11 15:42:39 2003

RESULT 14

US-09-978-295A-526
 ; Sequence 526, Application US/09978295A
 ; Patent No. US20020156006A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2630P1C11
 ; CURRENT APPLICATION NUMBER: US/09/978,295A
 ; CURRENT FILING DATE: 2001-10-15
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26

; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079786
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079920
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/079923
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/080105
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080107
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080194
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080327
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080328
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080333
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080334
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/081070
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081049
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081071
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081195
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081203
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081229
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081955
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081817
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081819
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081952
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081838
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/082568
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082569
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082704
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082804
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082700
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082797
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082796
 ; PRIOR FILING DATE: 1998-04-23
 ; PRIOR APPLICATION NUMBER: 60/083336
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/083392
 ; PRIOR FILING DATE: 1998-04-29
 ; PRIOR APPLICATION NUMBER: 60/083495

```

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 61.1%; Score 33; DB 9; Length 736;
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 EEVVPXGMDY 10
Db 331 EPVVYVGMDY 340

```

RESULT 15

```

; US-09-978-697-526
; Sequence 526, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978.697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728

```


APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500

us-09-909-164-13.rapb

Wed Jun 11 15:42:39 2003

;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGNDY 10
Db 331 EPVVYGM DY 340

RESULT 17
US-09-999-832A-526
Sequence 526, Application US/09999832A
Publication No. US20020192706A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fond, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher

;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Kuo, Sophia S.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James;
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Shelton, David L.
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2630PLC63
;; CURRENT FILING DATE: 2001-10-24
;; PRIOR APPLICATION NUMBER: 09/918585
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194

; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1% Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGMDY 340

RESULT 18

US-09-978-189-526
; Sequence 526, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napler, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C7
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US/09/978,189
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640

; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084598
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084627
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084643
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/085339
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085338
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085323
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085582
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085700
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085689
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085579
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085580
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085573
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085704
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 Db 331 EPVVYGYMDY 340

RESULT 19

US-10-174-590-420
 ; Sequence 420, Application US/10174590
 ; Publication No. US20030008352A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C42
 ; CURRENT APPLICATION NUMBER: US/10/174,590
 ; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 420
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-174-590-420

Query Match 61.1%; Score 33; DB 9; Length 736;
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 Db 331 EPVVYGYMDY 340

RESULT 20

US-10-176-758-420
 ; Sequence 420, Application US/10176758
 ; Publication No. US20030008353A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C104
 ; CURRENT APPLICATION NUMBER: US/10/176,758
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 420
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-176-758-420

Query Match 61.1%; Score 33; DB 9; Length 736;
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 Db 331 EPVVYGYMDY 340

RESULT 21

US-10-175-737-420
 ; Sequence 420, Application US/10175737
 ; Publication No. US20030013153A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C50
 ; CURRENT APPLICATION NUMBER: US/10/175,737
 ; CURRENT FILING DATE: 2002-06-19
 ; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 420
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-175-737-420

Query Match 61.1%; Score 33; DB 9; Length 736;

Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 0; Indels 3; Gaps 0;

Qy 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGM DY 340

RESULT 22

US-10-173-706-420
; Sequence 420, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-420

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 0; Indels 3; Gaps 0;

Qy 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGM DY 340

RESULT 23

US-10-175-738-420
; Sequence 420, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-175-738-420

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 0; Indels 3; Gaps 0;

Qy 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGM DY 340

RESULT 24

US-10-175-752-420
; Sequence 420, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-420

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 0; Indels 3; Gaps 0;

Qy 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGM DY 340

RESULT 25

US-10-176-482-420
; Sequence 420, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420

; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-420

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVYGYMDY 340

Search completed: June 10, 2003, 14:35:44
Job time : 16.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds
(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 EFWVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:*

- 1: pir1:**
- 2: pir2:**
- 3: pir3:**
- 4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40	74.1	156	S54619	hypothetical prote
2	38	70.4	363	D69551	conserved hypothet
3	36	66.7	102	A42452	V1 protein - tobac
4	36	66.7	341	S72649	sucrose-phosphate
5	36	66.7	348	S72650	sucrose-phosphate
6	36	66.7	460	G96764	unknown protein F2
7	36	66.7	1049	J04783	sucrose-phosphate
8	36	66.7	1068	J01329	sucrose-phosphate
9	36	66.7	1081	T09837	sucrose-phosphate
10	36	66.7	1083	T04062	sucrose-phosphate
11	36	66.7	1084	T04103	sucrose-phosphate
12	35	64.8	425	T24111	hypothetical prote
13	35	64.8	433	H87660	peptidoglycan-bind
14	35	64.8	440	H2784	probable alkaline
15	35	64.8	1150	T20173	hypothetical prote
16	35	64.8	1474	F69009	probable membrane
17	35	64.8	2747	B49132	fat facets (faf) s
18	34	63.0	99	S00210	plastocyanin b - L
19	34	63.0	155	S38255	plastocyanin precu
20	34	63.0	168	S38208	plastocyanin b pre
21	34	63.0	290	D38182	O6-methylguanidine-D
22	34	63.0	290	AG3104	6-O-methylguanidine-
23	34	63.0	296	F72745	hypothetical prote
24	34	63.0	357	G69290	probable hexosyltr
25	34	63.0	366	G69350	L-lactate dehydrog
26	34	63.0	565	E86665	ABC transporter AT
27	34	63.0	566	A70164	phenylalanine-trna
28	34	63.0	587	F81138	succinate dehydrog
29	34	63.0	906	T48898	disease resistance

30	34	63.0	908	2	T48899	disease resistance
31	34	63.0	1062	2	F83335	RND multidrug effl
32	34	63.0	1062	2	T30830	hypothetical prote
33	34	63.0	3472	2	T31308	hypothetical 367K
34	33	61.1	97	2	A99427	partial transposas
35	33	61.1	128	2	A90471	hypothetical prote
36	33	61.1	172	2	S27021	fibroblast growth
37	33	61.1	184	2	E90335	hypothetical prote
38	33	61.1	225	2	S57810	hypothetical prote
39	33	61.1	247	2	A96001	conserved hypothet
40	33	61.1	257	2	A95456	unknown protein [i
41	33	61.1	262	2	F90298	transposase ISC105
42	33	61.1	267	2	C90307	transposase ISC105
43	33	61.1	276	2	C64417	hypothetical prote
44	33	61.1	283	2	G83055	pantoate-beta-alan
45	33	61.1	299	2	E90487	transposase ISC105
46	33	61.1	299	2	H90352	transposase ISC105
47	33	61.1	307	2	F84330	hypothetical prote
48	33	61.1	394	2	F82491	ferri-siderophore r
49	33	61.1	421	1	DETCM	acyl-CoA dehydroge
50	33	61.1	670	2	S22293	zinc finger protei
51	33	61.1	797	2	S38579	fibroblast growth
52	33	61.1	800	1	TVH02F	fibroblast growth
53	33	61.1	800	2	A48991	heparin-binding gr
54	33	61.1	801	2	I55363	fibroblast growth
55	33	61.1	806	2	A35963	protein-tyrosine k
56	33	61.1	840	2	AG0526	penicillin-binding
57	33	61.1	840	2	T39116	probable sulfate p
58	33	61.1	846	2	S57580	penicillin-binding
59	33	61.1	877	2	T40413	sulfate permease -
60	33	61.1	982	1	VCLJLK	env polyprotein -
61	33	61.1	1064	2	F86182	hypothetical prote
62	33	61.1	1401	2	G82336	DNA-directed RNA p
63	33	61.1	2717	2	A34203	DNA-binding protei
64	33	61.1	2831	2	T31419	cyclic beta 1-2 gl
65	33	61.1	2867	2	AG3481	cellobiose-phospho
66	32.5	60.2	472	1	B53236	transcription fact
67	32	59.3	97	2	JW0011	plastocyanin - car
68	32	59.3	165	2	AG1272	thiol peroxidases
69	32	59.3	165	2	AH1635	thiol peroxidases
70	32	59.3	175	2	S36749	transcription fact
71	32	59.3	180	2	AG0504	fimbrial chain fim
72	32	59.3	231	1	I5ECP4	L-ribulose-phospha
73	32	59.3	231	2	A90637	L-ribulose-5-phosp
74	32	59.3	231	2	A85488	L-ribulose-5-phosp
75	32	59.3	231	2	AB0515	L-ribulose-5-phosp

ALIGNMENTS

RESULT 1
S54619
Hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O2612; Hypothetical protein YOL303.3
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
C;Accession: S54619; S66879
R;de Haan, M.; Maarse, A.C.; Grivell, L.A.
Submitted to the EMBL Data Library, May 1995
A;Reference number: S54617
A;Accession: S54619
A;Molecule type: DNA
A;Residues: 1-156 <DEH>
A;Cross-references: EMBL:X87331; NID:gl041652; PIDN:CAA60762.1; PID:g829123
R;de Haan, M.; Grivell, L.A.; Maarse, A.C.
Submitted to the Protein Sequence Database, July 1996
A;Reference number: S66877
A;Accession: S66879
A;Molecule type: DNA
A;Residues: 1-156 <DEH>
A;Cross-references: EMBL:Z74920; NID:gl420109; PIDN:CAA99201.1; PID:gl420111; MIPS
A;Experimental source: strain S288C

C:Genetics:
A:Cross-references: SGD:S0005539
A:Map position: 15R
C:Superfamily: hypothetical protein YOR013w

Query Match 74.1%; Score 40; DB 2; Length 156;
Best Local Similarity 77.8%; Pred. No. 1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
||| ||||
DB 50 EVMPGLGMDY 58

RESULT 2

D69551
conserved hypothetical protein AF2411 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69551
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Ventet, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: D69551
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-363 <KLE>

A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AAB91255.1; PID:g265068

Query Match 70.4%; Score 38; DB 2; Length 363;
Best Local Similarity 54.5%; Pred. No. 6.8;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
|:||||:|
DB 120 ENIVPYGIDFS 130

RESULT 3

A42452
V1 protein - tobacco yellow dwarf virus (strain Australia)
C:Species: tobacco yellow dwarf virus
C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
C:Accession: A42452
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
Virology 187, 633-642, 1992
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yell
A:Reference number: A42452; MUID:92188538; PMID:1546458
A:Accession: A42452
A:Molecule type: DNA
A:Residues: 1-102 <MOR>
A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 66.7%; Score 36; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 4.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDYS 11
:||||:|
DB 7 QVVPNGINYS 16

RESULT 4

S72649
sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Citrus unshiu (fragment)
C:Species: Citrus unshiu
C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C:Accession: S72649

R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
Mol. Gen. Genet. 252, 346-351, 1996
A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate
A:Reference number: S72648; MUID:96439842; PMID:8842155

A:Accession: S72649
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-341 <KOM>

A:Cross-references: EMBL:AB006319; NID:g2588891; PIDN:BAA23215.1; PID:g2588892
A:Experimental source: fruit, cv. Miyagawa-Wase
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1999
C:Genetics:
A:Gene: SPS2

C:Function:
A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fr
A:Pathway: sucrose biosynthesis
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homo
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:1-341/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SPSP>

Query Match 66.7%; Score 36; DB 2; Length 341;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||||
DB 228 VIPPGMDFS 236

RESULT 5

S72650
sucrose-phosphate synthase (EC 2.4.1.14) isoform 3 - Citrus unshiu (fragment)
C:Species: Citrus unshiu
C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C:Accession: S72650
R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
Mol. Gen. Genet. 252, 346-351, 1996
A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate
A:Reference number: S72648; MUID:96439842; PMID:8842155
A:Accession: S72650
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-348 <KOM>

A:Cross-references: EMBL:AB006660; NID:g2351059; PIDN:BAA22071.1; PID:g2351060
A:Experimental source: fruit, cv. Miyagawa-Wase
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997
C:Genetics:
A:Gene: SPS3
C:Function:

A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fr
A:Pathway: sucrose biosynthesis
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homo
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:1-348/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SPSP>

Query Match 66.7%; Score 36; DB 2; Length 348;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||||
DB 234 VIPPGMDFS 242

RESULT 6

G96764
unknown protein F25P22.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96764
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Al
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-460 <STO>
A:Cross-references: GB:AE005173; NID:g6692750; PIDN:AAF24856.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25P22.17
A:Map position: 1

Query Match 66.7%; Score 36; DB 2; Length 460;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVPXGMDY 10
DB 218 EEDVPSAMDY 227

RESULT 7
JC4783
sucrose-phosphate synthase (EC 2.4.1.14) - rice
C:Species: Oryza sativa (rice)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 18-Jun-1999
C:Accession: JC4783
R:Valdez-Alarcon, J.J.; Ferrando, M.; Salerno, G.; Jimenez-Morilla, B.; Herrera-Estrella
Gene 170, 217-222, 1996
A:Title: Characterization of a rice sucrose-phosphate synthase-encoding gene.
A:Reference number: JC4783; MUID:96235138; PMID:8666248
A:Accession: JC4783
A:Molecule type: mRNA
A:Residues: 1-1049 <VAL>
A:Cross-references: GB:033175; NID:g1449931; PIDN:RAC49379.1; PID:g988270
A:Note: UDPglucose-fructose-phosphate glucosyltransferase; Sucrosephosphate-UDPglucosylt
C:Comment: This enzyme catalyzes the formation of sucrose-phosphate form UDP-glucose and
C:Genetics:
A:Gene: Sps1
A:Introns: 24/1; 103/3; 183/3; 205/3; 435/3; 475/3; 519/3; 578/3; 596/3; 617/3; 931/3; 9
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc
A:Pathway: sucrose biosynthesis
A:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:180-663/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1049;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 436 VIPPGMDFS 444

RESULT 8
JQ1329
sucrose-phosphate synthase (EC 2.4.1.14) - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JQ1329; PQ0260
R:Worrell, A.C.; Bruneau, J.M.; Summerfelt, K.; Boersig, M.; Voelker, T.A.
Plant Cell 3, 1121-1130, 1991
A:Title: Expression of a maize sucrose phosphate synthase in tomato alters leaf carbohyd
A:Reference number: JQ1329; MUID:92338837; PMID:1840396
A:Accession: JQ1329
A:Molecule type: mRNA

A:Residues: 1-1068 <WOR>
A:Cross-references: GB:M97550; NID:g168625; PIDN:AAA33513.1; PID:g168626
A:Accession: PQ0260
A:Molecule type: protein
A:Residues: 71-74; 206-212; 471-481; 872-892 <WOR1>
C:Comment: This enzyme transfers the glucosyl group from UDPglucose to fructose-6-ph
C:Comment: This enzyme is involved in the regulation of carbon partitioning in the l
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-
A:Pathway: sucrose biosynthesis
A:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:178-666/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 1; Length 1068;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 435 VIPPGMDFS 443

RESULT 9
T09837
sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Craterostigma plantagineum
C:Species: Craterostigma plantagineum
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09837
R:Ingram, J.; Chandler, J.W.; Gallagher, L.; Salamini, F.; Bartels, D.
Plant Physiol. 115, 113-121, 1997
A:Title: Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to
A:Reference number: 126874; MUID:97451773; PMID:9306694
A:Accession: T09837
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1081 <INC>
A:Cross-references: EMBL:Y11795; NID:g2190349; PIDN:CAA72491.1; PID:g2190350
A:Experimental source: ABA-treated callus
C:Genetics:
A:Gene: sps2
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and I
A:Pathway: sucrose biosynthesis
A:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:176-674/Domain: sucrose/sucrose-phosphate synthase homology <SSS>

Query Match 66.7%; Score 36; DB 2; Length 1081;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 445 VIPPGMDFS 453

RESULT 10
T04062
sucrose-phosphate synthase homolog F28M11.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 16-Jul-1999
C:Accession: T04062
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15184
A:Accession: T04062
A:Molecule type: DNA
A:Residues: 1-1083 <BEV>
A:Cross-references: EMBL:AL049487
A:Experimental source: cultivar Columbia; BAC clone F28M11
C:Genetics:
A:Map position: 4

A;Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-1150 <WIL>
A:Cross-references: EMBL:Z81486; PIDN:CAB03994.1; GSPDB:GN00023; CESP:C53A5.2
A:Experimental source: clone C53A5
R:Matthews, L.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19608
A:Accession: T23857
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1150 <WIL2>
A:Cross-references: EMBL:Z78015; PIDN:CAB01437.1; GSPDB:GN00023; CESP:C53A5.2
A:Experimental source: clone R02D5
C:Genetics:
A:Gene: CESP:C53A5.2
A:Map position: 5
A:Introns: 33/3; 63/3; 132/3; 169/3; 221/3; 299/3; 379/2; 423/2; 438/2; 471/1; 513/2; 65/2
Query Match 64.8%; Score 35; DB 2; Length 1150;
Best Local Similarity 66.7%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 1

QY 3 VVPXGMDYS 11
DB 562 VLPVGIDYS 570
|:|:|:|:|:|

RESULT 16
F69009
probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69009
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Olu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69009
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1474 <MTH>
A:Cross-references: GB:AE000878; GB:AE000666; NID:g2622171; PIDN:AAB85563.1; PID:g262217
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1074
C:Keywords: duplication

Query Match 64.8%; Score 35; DB 2; Length 1474;
Best Local Similarity 55.6%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 3; Mismatches 1

QY 2 EVVPXGMDY 10
DB 1238 DVLPLAGLDY 1246
|:|:|:|:|

RESULT 17
B49132
fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
C:Accession: B49132; A49132
R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
Development 116, 985-1000, 1992
A:Title: The fat facets gene is required for Drosophila eye and embryo development.
A:Reference number: A49132; MUID:93202020; PMID:1295747
A:Contents: isogenic st
A:Accession: B49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2747 <FIS>

A:Cross-references: GB:L04959; NID:gl57411; PIDN:AAF01345.1; PID:g6013474
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBI:129008, NCBIP:12
A:Accession: A49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2704, 'VT', 2707, 'ANNV' <FI2>
A:Cross-references: GB:L04958; NID:gl57410; PIDN:AAF01346.1; PID:g6013475
A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:129008, NCBIP:12
C:Keywords: alternative splicing

Query Match 64.8%; Score 35; DB 2; Length 2747;
Best Local Similarity 54.5%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 3

QY 1 EEVVPXGMDYS 11
DB 1394 EVIVPDQDFS 1404
|:|:|:|:|:|

RESULT 18
S00210
plastocyanin b - Lombardy poplar
C:Species: Populus nigra var. italica (Lombardy poplar)
C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 20-Apr-2000
C:Accession: S00210
R:Dimitrov, M.I.; Egorov, C.A.; Donchev, A.A.; Atanasov, B.P.
FEBS Lett. 226, 17-22, 1987
A:Title: Complete amino acid sequence of poplar plastocyanin b.
A:Reference number: S00210
A:Accession: S00210
A:Molecule type: protein
A:Residues: 1-99 <DIM>
C:Superfamily: Plastocyanin
C:Keywords: Chloroplast; copper; electron transfer; metalloprotein
F:37,84,87,92/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 63.0%; Score 34; DB 2; Length 99;
Best Local Similarity 54.5%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
DB 43 EDAPVSGVDVS 53
|:|:|:|:|

RESULT 19
S38255
plastocyanin precursor - barley
C:Species: Hordeum vulgare (barley)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jun-1999
C:Accession: S38255; S00206
R:Nielsen, P.S.; Gausing, K.
Eur. J. Biochem. 217, 97-104, 1993
A:Title: In vitro binding of nuclear proteins to the barley plastocyanin gene prom
A:Reference number: S38255; MUID:94039081; PMID:8223592
A:Accession: S38255
A:Molecule type: DNA
A:Residues: 1-155 <NIE1>
A:Cross-references: EMBL:Z28347; NID:g431919; PIDN:CAA82201.1; PID:g431920
A:Experimental source: strain NK 1558
R:Nielsen, P.S.; Gausing, K.
FEBS Lett. 225, 159-162, 1987
A:Title: The precursor of barley plastocyanin: sequence of cDNA clones and gene ex
A:Reference number: S00206
A:Accession: S00206
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-119, 'T', 121-155 <NIE2>
A:Cross-references: EMBL:Y00704; NID:g22704
A:Note: not compared to nucleotide translation
C:Genetics:
A:Genome: nuclear

C;Function:
A;Description: mediates the transfer of electrons from cytochrome b6/f to photosystem I
C;Superfamily: plastocyanin
C;Keywords: chloroplast; copper; electron transfer; metalloprotein
F;1-58/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;59-155/Product: plastocyanin #status predicted <MAT>
F;95,140,143,148/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 63.0%; Score 34; DB 2; Length 155;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
: : : : :
Db 101 EDAPVSGVDVS 111

RESULT 20
S38208
plastocyanin b precursor - black poplar
C;Species: Populus nigra (black poplar)
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-May-2000
C;Accession: S58208
R;Reichert, J.; Jenzelewski, V.; Haehnel, W.
submitted to the EMBL Data Library, July 1995
A;Description: Kinetic studies of recombinant poplar plastocyanins.
A;Reference number: S58208
A;Accession: S58208
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-168 <REI>
C;Cross-references: EMBL:Z50186; NID:g929814; PIDN:CAA90565.1; PID:g929815
C;Superfamily: plastocyanin
C;Keywords: copper; electron transfer; metalloprotein
F;106,153,156,161/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 63.0%; Score 34; DB 2; Length 168;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
: : : : :
Db 112 EDAPVSGVDVS 122

RESULT 21
D98182
O6-methylguanine-DNA methyltransferase PA2118 [imported] - Agrobacterium tumefaciens (st
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: D98182
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: D98182
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88982.1; PID:gl5158766; GSPDB:GN00170
A;Gene: AGR_L_818
A;Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
: : : : :
Db 9 EDITPIGSDY 18

RESULT 22

AG3104

6-O-methylguanine-DNA methyltransferase [imported] - Agrobacterium tumefaciens (str
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AG3104
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; W
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; M
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AG3104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45253.1; PID:gl7742937; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: ada
A;Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;

Best Local Similarity 50.0%; Pred. No. 35;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10

: : : : :

Db 9 EDITPIGSDY 18

RESULT 23

F72745

hypothetical protein APE0493 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C;Accession: F72745

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.;

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, T.

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: F72745

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-296 <KAW>

A;Cross-references: DDBJ:AP000059; NID:g5103911; PIDN:BAA79458.1; PID:dl043244; PID:

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0493

C;Superfamily: Aeropyrum pernix hypothetical protein APE0493

Query Match 63.0%; Score 34; DB 2; Length 296;

Best Local Similarity 50.0%; Pred. No. 36;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMDYS 11

: : : : :

Db 3 ETLPGGLDYT 12

RESULT 24

G69290

probable hexosyltransferase (EC 2.4.1.-) AF0327 [similarity] - Archaeoglobus fulgidu

C;Species: Archaeoglobus fulgidus

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C;Accession: G69290

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Wed Jun 11 15:42:45 2003

us-09-909-164-13.rpr

Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69290
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-357 <KLE>
 A:Cross-references: GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AAB90909.1; PID:g265031
 C:Superfamily: probable hexosyltransferase ytxN
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 63.0%; Score 34; DB 1; Length 357;
 Best Local Similarity 55.6%; Pred. No. 44;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

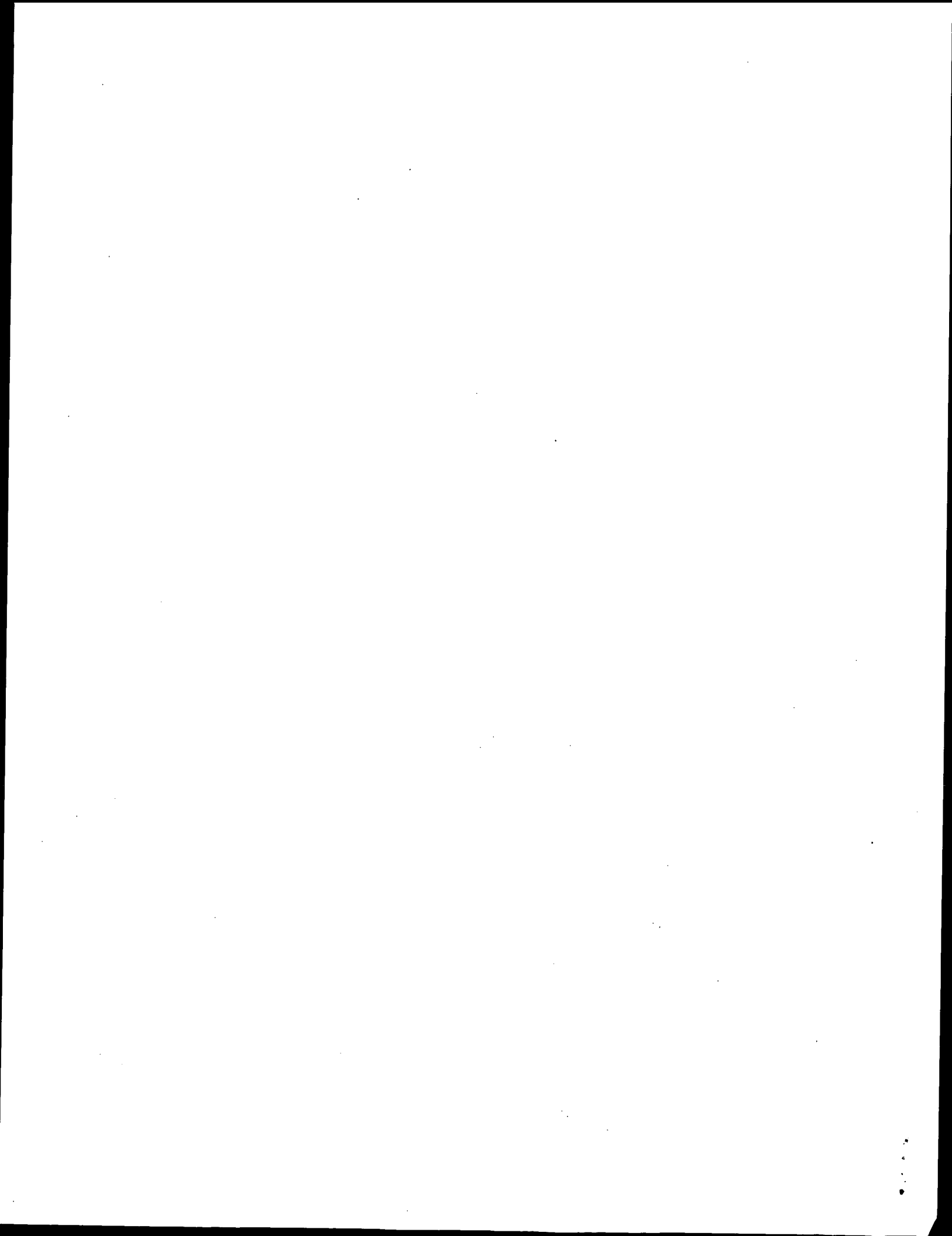
QY 2 EVVPXGMDY 10
 ||:| |:
 DB 170 EVIPNGIDF 178

RESULT 25
 G69350
 L-lactate dehydrogenase, cytochrome-type (lldD) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence-revision 05-Dec-1997 #text-change 17-Mar-2000
 C:Accession: G69350
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69350
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-366 <KLE>
 A:Cross-references: GB:AE001049; GB:AE000782; NID:g2689372; PIDN:AAB90435.1; PID:g264980
 C:Superfamily: (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology

Query Match 63.0%; Score 34; DB 2; Length 366;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9
 |::|| |:
 DB 289 EKVPTGVD 297

Search completed: June 10, 2003, 13:49:15
 Job time : 12.2143 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 EEVVPXGMDYS I1

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	70.4	1058	1	CARB_FUSNN
2	36	66.7	102	1	Q8r986 fusobacteri
3	36	66.7	1049	1	P31619 tobacco yel
4	36	66.7	1068	1	Q43802 oryza sativ
5	36	66.7	1081	1	P31927 zea mays (m
6	35	64.8	2747	1	O04933 craterostig
7	34.5	63.9	748	1	P55824 drosophila
8	34	63.0	154	1	Q9nr64 homo sapien
9	34	63.0	155	1	P20423 oryza sativ
10	34	63.0	168	1	P08248 hordeum vul
11	34	63.0	566	1	P11970 populus nig
12	33	61.1	276	1	P94283 borrelia bu
13	33	61.1	283	1	Q58349 methanococc
14	33	61.1	394	1	Q9nmv9 pseudomonas
15	33	61.1	421	1	Q9kmy3 vibrio chol
16	33	61.1	421	1	P08503 rattus norv
17	33	61.1	423	1	O52250 halomonas e
18	33	61.1	787	1	Q9zeu7 halomonas e
19	33	61.1	801	1	O60344 homo sapien
20	33	61.1	806	1	Q61851 mus musculu
21	33	61.1	877	1	Q18460 gallus gall
22	33	61.1	982	1	O74377 schizosacch
23	33	61.1	1401	1	P27399 simian foam
24	33	61.1	2717	1	Q9kv29 vibrio chol
25	32.5	60.2	472	1	P15922 homo sapien
26	32	59.3	97	1	P20422 daucus caro
27	32	59.3	175	1	P19102 xenopus lae
28	32	59.3	231	1	Q04667 rattus norv
29	32	59.3	231	1	P08203 escherichia
30	32	59.3	231	1	P06190 salmonella
31	32	59.3	233	1	Q9wrl1 thermotoga
32	32	59.3	288	1	Q04827 rattus norv
33	32	59.3	289	1	P30279 homo sapien
					P30280 mus musculu

34	32	59.3	291	1	CGD1_BRARE	Q90459 brachydanio
35	32	59.3	291	1	CGD1_XENLA	P50755 xenopus lae
36	32	59.3	291	1	CGD2_CHICK	P49706 gallus gall
37	32	59.3	291	1	CGD2_XENLA	P53782 xenopus lae
38	32	59.3	292	1	CGD1_CHICK	P55169 gallus gall
39	32	59.3	292	1	CGD3_HUMAN	P30281 homo sapien
40	32	59.3	295	1	CGD1_HUMAN	P24385 homo sapien
41	32	59.3	295	1	CGD1_MOUSE	P25322 mus musculu
42	32	59.3	295	1	CGD1_RAT	P39948 rattus norv
43	32	59.3	338	1	MTBA_METBA	Q30640 methanosarc
44	32	59.3	472	1	ET2A_XENLA	Q91712 xenopus lae
45	32	59.3	561	1	HNFB_XENLA	Q91910 xenopus lae
46	32	59.3	759	1	SC11_YEAST	P32784 saccharomyc
47	32	59.3	866	1	RECE_ECOLI	P15032 escherichia
48	32	59.3	995	1	HIP1_HUMAN	O00291 homo sapien
49	32	59.3	1176	1	NIR_NEUCR	P38681 neurospora
50	32	59.3	1258	1	ACNL_MOUSE	Q61137 mus musculu
51	32	59.3	1394	1	LTBS_HUMAN	P22054 homo sapien
52	32	59.3	1498	1	Y1A9_CLOAB	Q04351 clostridium
53	32	59.3	1595	1	LTBL_HUMAN	Q14766 homo sapien
54	32	59.3	1712	1	LTBL_RAT	Q00918 rattus norv
55	32	59.3	3174	1	CHAC_HUMAN	Q96r17 homo sapien
56	31	57.4	98	1	PLAS_ENTPR	P07465 enteromorph
57	31	57.4	98	1	PLAS_ULVPE	P13133 ulva arasak
58	31	57.4	98	1	PLAS_TUBOB	P56274 ulva pertus
59	31	57.4	99	1	PLAS_TUBOB	P00298 rumex obtus
60	31	57.4	99	1	REV_SIVC2	P35476 nicotiana t
61	31	57.4	124	1	HES3_MOUSE	P17280 chimpanzee
62	31	57.4	175	1	YK14_CABEL	Q61657 mus musculu
63	31	57.4	258	1	YK10_METJA	O58607 methanococc
64	31	57.4	319	1	YHAI_CRYPA	P10941 cryptonectr
65	31	57.4	327	1	YK14_CABEL	P34338 caenorhabdi
66	31	57.4	338	1	MTBA_METAC	P38869 methanosarc
67	31	57.4	346	1	HYPE_BRAJA	P31906 bradyrhizob
68	31	57.4	391	1	LE11_METTH	Q37667 methanobact
69	31	57.4	427	1	TOLE_HAEIN	P44677 haemophilus
70	31	57.4	469	1	LE11_KLOLA	P53998 kluyveromyc
71	31	57.4	529	1	ENP3_HUMAN	O75355 homo sapien
72	31	57.4	529	1	GUAA_MYCLE	P46810 mycobacteri
73	31	57.4	625	1	GIDA_STAAM	Q99qt4 staphylococ
74	31	57.4	692	1	DNK1_SYNV3	Q55154 synechocyst
75	31	57.4	788	1	CY14_NEUCR	P23622 neurospora

ALIGNMENTS

RESULT 1
CARB_FUSNN STANDARD: PRT: 1058 AA.
ID CARB_FUSNN
AC Q8RG86;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
GN CARB OR FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykdis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusck G., Haselkorn R.,
RA Fontaine M., Kyripides N., Overbeek R.,
RA "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
RT J. Bacteriol. 184:2005-2018(2002).
RL -1- CATALYTIC ACTIVITY: 2 ATP + H(2)O = 2 ADP +

CC phosphate + L-glutamate + carbamoyl phosphate.
 CC -|- COPACTOR: Binds three manganese ions (By similarity).
 CC -|- PATHWAY: Arginine biosynthesis.
 CC -|- PATHWAY: Pyrimidine biosynthesis; first step.
 CC -|- SUBUNIT: Composed of two chains; the small (or glutamine) chain
 CC promotes the hydrolysis of glutamine to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate. (By
 CC similarity).
 CC -|- SIMILARITY: BELONGS TO THE CARB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE010554; AAL04625.1; ALT_INIT.
 CC InterPro; IPR005483; CPase_L.
 CC InterPro; IPR005479; CPase_L_D2.
 CC InterPro; IPR005480; CPase_L_D3.
 CC InterPro; IPR005481; CPase_L_N.
 CC InterPro; IPR004362; MGS_like.
 CC Pfam; PF00289; CPase_L_chain; 2.
 CC Pfam; PF02786; CPase_L_D2; 2.
 CC Pfam; PF02787; CPase_L_D3; 1.
 CC Pfam; PF02142; MGS; 1.
 CC PRINTS; PR00098; CPASE.
 CC PROSITE; PS00866; CPASE_1; 2.
 CC PROSITE; PS00867; CPASE_2; 2.
 CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 CC ATP-binding; Manganese; Complete proteome.
 CC DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 CC DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 CC DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 CC DOMAIN 930 1058 ALLOSTERIC DOMAIN.
 CC REPEAT 1 546
 CC REPEAT 547 1058
 CC NP_BIND 153 210
 CC NP_BIND 302 352
 CC METAL 284 284
 CC METAL 298 298
 CC METAL 300 300
 CC METAL 820 820
 CC METAL 832 832
 CC SEQUENCE 1058 AA; 117451 MW; ED/037AF77C1E39F CRC64;
 CC
 CC Query Match 70.4%; Score 38; DB 1; Length 1058;
 CC Best Local Similarity 60.0%; Pred. No. 7.8;
 CC Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 EVVPXGMDYS 11
 CC Db 190 EIVPGLNYS 199
 CC
 CC RESULT 2
 CC Y1LK TYDVA
 CC ID Y1LK TYDVA STANDARD; PRT; 102 AA.
 CC AC P31619;
 CC DT 01-JUL-1993 (Rel. 26, Created)
 CC DT 01-JUL-1993 (Rel. 26, Last sequence update)
 CC DT 01-OCT-1993 (Rel. 27, Last annotation update)
 CC DE Hypothetical 11.2 kDa protein.
 CC GN V1.
 CC OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
 CC OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 CC OX NCBI_TaxID=31599;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=92188538; PubMed=1546458;
 CC RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT "The nucleotide sequence of the infectious cloned DNA component of
 RT tobacco yellow dwarf virus reveals features of geminiviruses
 RT infecting monocotyledonous plants.";
 RT Virology 187:633-642(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81103; AAA47947.1; -
 CC PIR; A42452; A42452.
 CC InterPro; IPR002621; Gemini_mov.
 CC Pfam; PF01708; Gemini_mov; 1.
 CC KW Hypothetical protein.
 CC SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;
 CC
 CC Query Match 66.7%; Score 36; DB 1; Length 102;
 CC Best Local Similarity 60.0%; Pred. No. 1.7;
 CC Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 EVVPXGMDYS 11
 CC Db 7 QVVPGLNYS 16
 CC
 CC RESULT 3
 CC SPS_ORYSA
 CC ID SPS_ORYSA STANDARD; PRT; 1049 AA.
 CC AC Q43802;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE Sucrose-phosphate synthase (EC 2.4.1.14)
 CC DE (UDP-glucose-fructose-phosphate glucosyltransferase).
 CC OS Oryza sativa (Rice).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC OC Ehrhartoideae; Oryzaceae; Oryza.
 CC OX NCBI_TaxID=4530;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Indica-IR36; TISSUE=Leaf;
 CC RX MEDLINE=96235138; PubMed=8666248;
 CC RA Valdez-Alarcon J.J., Ferrando M., Jimenez-Morilla B.,
 CC RA Herrera-Estrella L.;
 CC RT "Characterization of a rice sucrose-phosphate synthase-encoding
 CC gene.";
 CC RL Gene 170:217-222(1996).
 CC -|- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
 CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
 CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
 CC PHOTOASSIMILATES OUT OF THE LEAF.
 CC -|- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate -> UDP +
 CC sucrose 6-phosphate.
 CC -|- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
 CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
 CC -|- PATHWAY: Sucrose synthesis.
 CC -|- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
 CC -|- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
 CC ENZYME FUNCTION (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; U33175; AAC49379.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Transferrase; Glycosyltransferase; Phosphorylation.
FT DOMAIN 22 29 POLY-GLY.
FT DOMAIN 695 698 POLY-GLU.
FT DOMAIN 775 779 POLY-ARG.
SQ SEQUENCE 1049 AA; 116455 MW; ED862E2819AA4B04 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1049;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGMDYS 11
DB 436 VIPPGMDFS 444

RESULT 4
SPS_MAIZE STANDARD; PRT; 1068 AA.
ID SPS_MAIZE
AC P31927;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase).
GN SPS.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 71-74; 206-212; 471-481 AND
RP 872-892
RC STRAIN=CV. PIONEER 3184; TISSUE=Leaf;
RA MEDLINE=92338837; PubMed=1840396;
RX Worrell A.C., Bruneau J.-M., Summerfelt K., Boersig M., Voelker T.A.;
RT "Expression of a maize sucrose phosphate synthase in tomato alters
RT leaf carbohydrate partitioning.";
RL Plant Cell 3:1121-1130(1991).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -1- PATHWAY: Sucrose synthesis.
CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
CC -1- DEVELOPMENTAL STAGE: GERMINATING SEEDS OR MATURE LEAVES.
CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M97550; AAA33513.1; -
DR PIR; J01329; J01329.
DR MaizeDB; 25294; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Transferrase; Glycosyltransferase; Phosphorylation.
FT DOMAIN 25 31 POLY-GLY.
SQ SEQUENCE 25

SQ SEQUENCE 1068 AA; 118575 MW; 074679B5E9A1D282 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1068;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGMDYS 11
DB 435 VIPPGMDFS 443

RESULT 5
SPS2_CRAPL STANDARD; PRT; 1081 AA.
ID SPS2_CRAPL
AC Q04933;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase 2).
GN SPS2.
OS Craterostigma plantagineum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Torenieae;
OC Craterostigma.
OX NCBI_TaxID=4153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97451773; PubMed=9306694;
RA Ingram J., Chandler J.W., Gallagher L., Salamini F., Bartels D.;
RT "Analysis of cDNA clones encoding sucrose-phosphate synthase in
RT relation to sugar interconversions associated with dehydration in the
RT resurrection plant Craterostigma plantagineum Hochst.";
RL Plant Physiol. 115:113-121(1997).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -1- PATHWAY: Sucrose synthesis.
CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Y11795; CAA72491.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Transferrase; Glycosyltransferase; Phosphorylation; Multigene family.
FT DOMAIN 245 248 POLY-SER.
FT DOMAIN 256 264 POLY-GLU.
FT DOMAIN 787 790 POLY-ARG.
SQ SEQUENCE 1081 AA; 120933 MW; DD142DC2F1A72900 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1081;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGMDYS 11
DB 445 VIPPGMDFS 453

RESULT 6
FAF_DROME
ID FAF_DROME STANDARD; PRT; 2747 AA.
AC P55824;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ubiquitin C-terminal hydrolase FAF (EC 3.1.2.15)
DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific protease
FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
GN FAF.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93202020; PubMed=1295747;
RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;
RT "The fat facets gene is required for Drosophila eye and embryo
development.";
RL Development 116:985-1000(1992).
CC -1- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A
ROLE IN COMPOUND EYE ASSEMBLY AND OCOGENESIS RESPECTIVELY. IN THE
LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS
PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY
CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR
NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND
COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR
FUNCTION.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
ubiquitin + a thiol.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L04959; AAF01345.1; -;
CC EMBL; L04958; AAF01346.1; -;
CC MEROPS; C19.007; -;
CC FlyBase; Fggn0005632; faf.
CC InterPro; IPR001394; UCH-2.
CC Pfam; PF00442; UCH-1; 1.
CC Pfam; PF00443; UCH-2; 1.
CC PROSITE; PS00972; UCH_2_1; 1.
CC PROSITE; PS00973; UCH_2_2; 1.
CC PROSITE; PS00973; UCH_2_3; 1.
CC Ubl conjugation pathway; Hydrolase; Thiol protease;
KW Developmental protein; Vision; Alternative splicing.
FT ACT_SITE 1677 1677 BY SIMILARITY.
FT ACT_SITE 1978 1978 BY SIMILARITY.
FT ACT_SITE 1986 1986 BY SIMILARITY.
FT VARSPLOC 2705 2747 KCRVLIKKLVSKDEATTSATTAATVTTSPATAS
FT VARIANT 2725 2725 S -> T.
FT ORQQL -> VTRANNV (IN SHORT ISOFORM).
SQ SEQUENCE 2747 AA; 307954 MW; 1D97659F7A7B2ADE CRC64;
Query Match 64.8%; Score 35; DB 1; Length 2747;
Best Local Similarity 54.5%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDYS 11

Db 1394 EVIVPDGQDFS 1404
| : | | | : |
RESULT 7
KHL1_HUMAN
ID KHL1_HUMAN STANDARD; PRT; 748 AA.
AC Q9NR64; Q9NR65; Q9P238; Q9H4X4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kelch-like protein 1.
GN KHL1 OR KIAA1490.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347694; PubMed=10888605;
RA Koob M.D., Nemes J.P., Benzow K.A.;
RT "The SCAR transcript is an antisense RNA to a brain-specific
transcript encoding a novel actin-binding protein (KHL1).";
RL Hum. Mol. Genet. 9:1543-1551(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 7:143-150(2000).
RN [3]
RP SEQUENCE OF 179-409 FROM N.A.
RA Ray N.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN ORGANIZING THE ACTIN CYTOSKELETON OF
THE BRAIN CELLS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF252283; AAF81719.1; -;
CC EMBL; AF252279; AAF81716.1; -;
CC EMBL; AB040923; BAA96014.1; ALT_INIT.
CC EMBL; AL353736; CAC16128.1; -;
CC Genew; HGNC:6352; KHL1.
CC MIM; 605332; -;
CC InterPro; IPR000210; BTB_POZ.
CC InterPro; IPR001798; Kelch.
CC Pfam; PF00651; BTB; 1.
CC Pfam; PF01344; Kelch; 6.
CC PRINTS; PR00501; KELCHREPEAT.
CC SMART; SM00225; BTB; 1.
CC PROSITE; PS0097; BTB; 1.
CC Cytoskeleton; Actin-binding; Repeat.
KW DOMAIN 43 88 SER-RICH.
FT DOMAIN 212 279 BTB.
FT REPEAT 460 506 KELCH 1.
FT REPEAT 507 553 KELCH 2.
FT REPEAT 555 600 KELCH 3.
FT REPEAT 601 647 KELCH 4.
FT REPEAT 649 700 KELCH 5.

FT REPEAT 701 747 KELCH 6.
SQ SEQUENCE 748 AA; 82680 MW; C11C43D8282F9FF9 CRC64;
Query Match 63.9%; Score 34.5; DB 1; Length 748;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 EVVVPXGMDY 10
Db 127 EVVVP-GMDF 135
RESULT 8
PLAS_ORYSA STANDARD; PRT: 154 AA.
AC P20423; Q9SBB8;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plastocyanin, chloroplast precursor.
GN PETE.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ilpoom; TISSUE=Leaf;
RA Lee J.-S.;
RT "Molecular cloning and characterization of plastocyanin precursor in rice."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-154.
RC STRAIN=cv. Japonica;
RX MEDLINE=89386623; PubMed=2780537;
RA Yano H., Kamo M., Tsugita A., Aso K., Nozu Y.;
RT "The amino acid sequence of plastocyanin from rice (Oryza sativa, subspecies japonica)."
RL Protein Seq. Data Anal. 2:385-389(1989).
CC -!- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.
CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.
CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AF093636; AAC78108.1; -.
CC PIR; S06105; S06105.
CC PIR; J03052; J03052.
CC HSP; P00289; 2PCF.
CC InterPro: IPR000923; BlueCu_1.
CC InterPro: IPR001235; Copper_blue.
CC Pfam: PF00127; copper-bind; 1.
CC PRINTS; PR00156; COPPERBLUE.
CC ProDom; PD001235; Copper_blue; 1.
CC PROSITE; PS00196; COPPER_BLUE; 1.
CC Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
KW Transit peptide.
FT TRANSIT 1 57 CHLOROPLAST.
FT CHAIN 58 154 PLASTOCYANIN.
FT DOMAIN 58 154 PLASTOCYANIN-LIKE.
FT METAL 94 94 COPPER (BY SIMILARITY).
FT METAL 139 139 COPPER (BY SIMILARITY).
FT METAL 142 142 COPPER (BY SIMILARITY).

FT METAL 147 147 COPPER (BY SIMILARITY).
SQ SEQUENCE 154 AA; 15577 MW; E45725D5B5F400D CRC64;
Query Match 63.0%; Score 34; DB 1; Length 154;
Best Local Similarity 54.5%; Pred. No. 6.8; 3; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 3;
QY 1 EVVVPXGMDYS 11
Db 100 EDVPSGVDS 110
RESULT 9
PLAS_HORVU STANDARD; PRT: 155 AA.
AC P08248; 1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plastocyanin, chloroplast precursor.
GN PETE.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Boml;
RA Nielsen O.S., Gausling K.;
RT "The precursor of barley plastocyanin: sequence of cDNA clones and gene expression in different tissues."
RL FEBS Lett. 225:159-162(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. NK 1558;
RX MEDLINE=94039081; PubMed=8223592;
RA Nielsen P., Gausling K.;
RT "In vitro binding of nuclear proteins to the barley plastocyanin gene promoter region."
RL Eur. J. Biochem. 217:97-104(1993).
CC -!- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.
CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.
CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; Y00704; CAA68696.1; -.
CC PIR; Z28347; CAA82201.1; -.
CC PIR; S00206; S00206.
CC HSP; P00289; 2PCF.
CC InterPro: IPR000923; BlueCu_1.
CC InterPro: IPR001235; Copper_blue.
CC Pfam: PF00127; copper-bind; 1.
CC PRINTS; PR00156; COPPERBLUE.
CC ProDom; PD001235; Copper_blue; 1.
CC PROSITE; PS00196; COPPER_BLUE; 1.
CC Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
KW Transit peptide.
FT TRANSIT 1 58 CHLOROPLAST.
FT CHAIN 59 155 PLASTOCYANIN.
FT DOMAIN 59 155 PLASTOCYANIN-LIKE.
FT METAL 95 95 COPPER (BY SIMILARITY).
FT METAL 140 140 COPPER (BY SIMILARITY).
FT METAL 143 143 COPPER (BY SIMILARITY).

FT METAL 148 148 COPPER (BY SIMILARITY).
 FT VARIANT 120 120 T -> N (IN CV. NK 1558).
 SQ SEQUENCE 155 AA; 15709 MW; DAA7EABE3F6F4F91 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 155;
 Best Local Similarity 54.5%; Pred. No. 6.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: | | | | |
 Db 101 EDVPSGVDS 111

RESULT 10
 PLAT_POPNI STANDARD; PRT; 168 AA.
 AC PLAT_POPNI
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin B, chloroplast precursor.
 GN PTE.
 OS Populus nigra (Lombardy poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Italica; TISSUE=Leaf;
 RA Reichert J., Jenzelewski V., Haehnel W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 70-168.
 RC STRAIN-cv. Italica;
 RA Dimitrov M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;
 RT "Complete amino acid sequence of poplar plastocyanin B";
 RL FEBS Lett. 226:17-22(1987).
 CC -1- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF
 CC POPLAR PLASTOCYANINS A AND B.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; 250186; CAA90565.1; -
 PIR; S00210; S00210.
 DR HSP; P00299; IPI.
 DR InterPro; IPR000923; BlueCu.1.
 DR InterPro; IPR001235; Copper_blue.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD001235; Copper_blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 DR Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 TRANSIT peptide; Multigene family.
 FT TRANSIT 1 69 CHLOROPLAST.
 FT CHAIN 70 168 PLASTOCYANIN B.
 FT DOMAIN 70 168 PLASTOCYANIN-LIKE.
 FT METAL 106 106 COPPER.
 FT METAL 153 153 COPPER.
 FT METAL 156 156 COPPER.
 FT METAL 161 161 COPPER.
 SQ SEQUENCE 168 AA; 16981 MW; F20DAGEA2038AEEA CRC64;

Query Match 63.0%; Score 34; DB 1; Length 168;
 Best Local Similarity 54.5%; Pred. No. 7.5;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: | | | | |
 Db 112 EDVPSGVDS 122

RESULT 11
 SYFB_BORBU STANDARD; PRT; 566 AA.
 ID SYFB_BORBU
 AC P94283;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phenylalanyl-trna synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
 DE tRNA ligase beta chain) (PHERS).
 GN PHET OR BB0514.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RA Barbour A.G., Hinnebusch J.;
 RT "Phenylalanyl-trna synthetase genes (alpha and beta subunits) and
 RT thioredoxin reductase gene of Borrelia burgdorferi";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 RT burgdorferi";
 RL Nature 390:580-586(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
 CC diphosphate + L-phenylalanyl-tRNA(Phe).
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
 CC CHAIN FAMILY. SUBFAMILY 2.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; U82978; AAB41019.1; -
 DR EMBL; AE001153; AAC66870.1; -
 DR TIGR; BB0514; -
 DR InterPro; IPR005147; B5.
 DR InterPro; IPR004531; Phet_arch.
 DR Pfam; PF03484; B5; 1.
 DR TIGRFAMS; TIGR00471; phet_arch; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 566 AA; 65173 MW; 9D48CB5D6D3B74B CRC64;

Query Match 63.0%; Score 34; DB 1; Length 566;

```
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMDY 10
Db 169 VPFGMDY 175

RESULT 12
Y939.METJA
ID Y939.METJA STANDARD; PRT; 276 AA.
AC Q58349;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0939.
GN MJ0939.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RC STRAIN=TAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1038-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE004886; AAG08116.1;
DR InterPro: IPR003721; Pantoate_ligase.
DR Pfam: P02569; Pantoate_ligase; 1.
DR TIGR: TIGR00018; pncC; 1.
DR Pantothenate biosynthesis; Ligase; Complete proteome.
KW Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 283 AA; 30836 MW; C494949AB40E14E7 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9
Db 96 EEMYPDGMD 104

RESULT 14
HMPA_VIBCH STANDARD; PRT; 394 AA.
ID HMPA_VIBCH
AC Q9KM73;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
GN HMP OR VCA0183.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RC STRAIN=FROM N.A.
RX MEDLINE=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS; ONE IS A HEME-CONTAINING
CC OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN
CC FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN
CC FLAVOHEMOPROTEINS SUBFAMILY.
```

```
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMDY 10
Db 169 VPFGMDY 175

RESULT 12
Y939.METJA
ID Y939.METJA STANDARD; PRT; 276 AA.
AC Q58349;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0939.
GN MJ0939.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RC STRAIN=TAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1038-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67537; AAB98946.1;
DR TIGR: MJ0939;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 276 AA; 33454 MW; 97BD69D392BC8FDF CRC64;

Query Match 61.1%; Score 33; DB 1; Length 276;
Best Local Similarity 45.5%; Pred. No. 21;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
Db 141 EEIENGNEHS 151

RESULT 13
PANC_PSEAE STANDARD; PRT; 283 AA.
ID PANC_PSEAE
AC Q9HV69;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate-beta-alanine enzyme).
GN PANC OR PA4730.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
```


CC CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC CC EMBL: AF031489; AAC15882.1; -

CC CC HSPSP; P12995; 1QJ3

CC CC InterPro: IPR000954; Aminotran_3.

CC CC InterPro: IPR004637; Dat.

CC CC Pfam: PF00202; aminotran_3; 1.

CC CC TIGRFAMs: TIGR00709; dat; 1.

CC CC PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.

CC CC Transferrase; Aminotransferase; PYRIDOXAL PHOSPHATE (POTENTIAL).

CC CC BINDING 267 267

CC CC SEQUENCE 421 AA; 46166 MW; A4AZE21596E1E16C CRC64;

CC CC Query Match 61.1%; Score 33; DB 1; Length 421;

CC CC Best Local Similarity 58.3%; Pred. No. 32;

CC CC Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

CC CC QY 1 EVVV--PXGMDY 10

CC CC III: | | | |

CC CC DB 91 EEVILKPRGLDY 102

CC CC RESULT 17

CC CC ECBI_HALEL STANDARD; PRT; 423 AA.

CC AC Q9ZEU7;

CC DT 30-MAY-2000 (Rel. 39, Created)

CC DT 30-MAY-2000 (Rel. 39, Last sequence update)

CC DT 16-OCT-2001 (Rel. 40, Last annotation update)

CC DE Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46) (L-diaminobutyrate acid transaminase) (diaminobutyrate transaminase).

CC GN ECTB.

CC OS Halomonas elongata.

CC OC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae;

CC OC Halomonas.

CC OX NCBI_TaxID=2746;

CC RN [1]

CC RP SEQUENCE FROM N.A.

CC RC STRAIN=DSM 3043;

CC RX MEDLINE=99123891; PubMed=9924816;

CC RA Canovas D., Vargas C., Calderon M.I., Ventosa A., Nieto J.J.;

CC RT "Characterization of the genes for the biosynthesis of the compatible solute ectoine in the moderately halophilic bacterium Halomonas elongata DSM 3043."

CC RL Syst. Appl. Microbiol. 21:487-497(1998).

CC CC -1- CATALYTIC ACTIVITY: L-2,4-diaminobutanoate + pyruvate = L-aspartate 4-semialdehyde + L-alanine.

CC CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

CC CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF ECTOINE (1,4,5,6-TETRAHYDRO-2-METHYL-4-PYRIMIDINE CARBOXYLIC ACID).

CC CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC CC EMBL: AJ011103; CAA09484.1; -

CC CC HSPSP; P12995; 1QJ3.

DR InterPro: IPR000954; Aminotran_3.

DR InterPro: IPR004637; Dat.

DR Pfam: PF00202; aminotran_3; 1.

DR TIGRFAMs: TIGR00709; dat; 1.

DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.

DR Transferrase; Aminotransferase; PYRIDOXAL PHOSPHATE (POTENTIAL).

DR FT BINDING 267 267

DR FT SEQUENCE 423 AA; 46200 MW; 735C6BCF5A88288C CRC64;

DR Query Match 61.1%; Score 33; DB 1; Length 423;

DR Best Local Similarity 58.3%; Pred. No. 32;

DR Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

DR QY 1 EVVV--PXGMDY 10

DR DB 91 EEVILKPRGLDY 102

DR RESULT 18

DR EC22_HUMAN STANDARD; PRT; 787 AA.

DR ID EC22_HUMAN STANDARD; PRT; 787 AA.

DR AC O60344; Q96NX4; Q96NX3;

DR DT 30-MAY-2000 (Rel. 39, Created)

DR DT 15-JUN-2002 (Rel. 41, Last sequence update)

DR DT 15-JUN-2002 (Rel. 41, Last annotation update)

DR DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).

DR GN ECE2 OR KIAA0604.

DR OS Homo sapiens (Human).

DR OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

DR OC Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

DR OX NCBI_TaxID=9606;

DR RN [1]

DR RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).

DR RA Lorenzo M.-N., Khan R.Y., Wang Y., Tai S.C., Chan G.C., Cheung A.H., Marsden P.A.;

DR RT "Human endothelin converting enzyme-2 (ECE2): characterization of mRNA species and chromosomal localization."

DR RL Biochim. Biophys. Acta 0:0-0(2002).

DR RN [2]

DR RP SEQUENCE FROM N.A. (ISOFORM B).

DR RC TISSUE=Brain;

DR RX MEDLINE=98290545; PubMed=9628581;

DR RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

DR RT "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."

DR RL DNA Res. 5:31-39(1998).

DR CC -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1 (BY SIMILARITY).

CC CC -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-Tip-1-Val-22 bond in the precursor.

CC CC -1- SUBCELLULAR LOCATION: Type II membrane protein.

CC CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; ECE-2A (shown here), ECE-2B and ECE-2C; are produced by alternative splicing.

CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC CC EMBL: AF428263; AAL30386.1; -

CC CC EMBL: AF428264; AAL30387.1; -

CC CC EMBL: AF192531; AAG28399.1; -

CC CC EMBL: AB011176; BAA25530.1; -

CC CC HSPSP; P08473; IDMT.

CC CC MEROPS: M13.003.

CC CC InterPro: IPR000718; Peptidase_M13.

CC CC InterPro: IPR000130; Zn_Mtpeptdse.

FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 305 352 TAGANTDKLEVLISLHNVTFEDAGEYTCLAGNSIGFSHRS
FT AWLVLP -> SWISENVEADARLRLANVSDRGGEYLCLRA
FT TNFIVAEKAFWLRVHGPOA (IN ISOFORM 2).
FT P -> L (IN REF. 2).
FT MISSING (IN REF. 2).
FT SEQUENCE 801 AA; 87758 MW; 68BC110212691705 CRC64;
Query Match 61.1%; Score 33; DB 1; Length 801;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PXGMDYS 11
DB 566 PPGMDYS 572
RESULT 20
CEK2_CHICK STANDARD; PRT; 806 AA.
AC P18460;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine kinase receptor CEK2 precursor (EC 2.7.1.112).
GN Gallus (Chicken).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90332672; PubMed=2165604;
RA Pasquale E.B.;
RT "A distinctive family of embryonic protein-tyrosine kinase
receptors";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: M35195; AAA48664.1; -;
DR HSP: A35963; A35963.
DR HSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig_3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00408; IGC2; 3.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Repeat; Signal.

FT SIGNAL 1 19
FT CHAIN 20 806
FT DOMAIN 20 806
FT TRANSMEM 365 389
FT DOMAIN 390 806
FT DOMAIN 54 114
FT DOMAIN 163 229
FT DOMAIN 262 340
FT DOMAIN 131 141
FT DOMAIN 466 755
FT NP_BIND 472 480
FT BINDING 502 502
FT ACT_SITE 611 611
FT MOD_RES 642 642
FT DISULFID 61 107
FT DISULFID 170 222
FT DISULFID 269 333
FT CARBOHYD 83 83
FT CARBOHYD 96 96
FT CARBOHYD 118 118
FT CARBOHYD 219 219
FT CARBOHYD 256 256
FT CARBOHYD 288 288
FT CARBOHYD 309 309
FT CARBOHYD 322 322
SQ SEQUENCE 806 AA; 89730 MW; B38B3C6D5F2314B6 CRC64;
Query Match 61.1%; Score 33; DB 1; Length 806;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 PXGMDYS 11
DB 566 PPGMDYS 572
RESULT 21
SULH_SCHPO STANDARD; PRT; 877 AA.
AC 074377;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable sulfate permease C3H7.02.
GN SPBC3H7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=q972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL031261; CAA20298.1; -
CC InterPro; IPR002645; STAS.
CC Pfam; PF00916; Sulfate_transp; 1.
CC TIGRFAMs; TIGR00815; sulp; 1.
CC PROSITE; PS01130; SLC26A; 1.
CC PROSITE; PS0801; STAS; 1.
CC Transmembrane.
CC TRANSMEM 133 153 POTENTIAL.
CC TRANSMEM 161 181 POTENTIAL.
CC TRANSMEM 186 206 POTENTIAL.
CC TRANSMEM 221 241 POTENTIAL.
CC TRANSMEM 243 263 POTENTIAL.
CC TRANSMEM 292 312 POTENTIAL.
CC TRANSMEM 329 349 POTENTIAL.
CC TRANSMEM 384 404 POTENTIAL.
CC TRANSMEM 424 444 POTENTIAL.
CC TRANSMEM 461 481 POTENTIAL.
CC TRANSMEM 484 504 POTENTIAL.
CC TRANSMEM 518 538 POTENTIAL.
CC TRANSMEM 543 563 POTENTIAL.
CC DOMAIN 594 747 STAS.
CC SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;

CC Query Match 61.1%; Score 33; DB 1; Length 877;
CC Best Local Similarity 66.7%; Pred. No. 70;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CC QY 3 VVPXGMDYS 11
CC |||||
CC Db 148 VVPQMSYA 156

CC RESULT 22
CC ENV_SFV3L
CC ID ENV_SFV3L STANDARD; PRT; 982 AA.
CC AC P27399;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC DE 15-JUL-1998 (Rel. 36, Last annotation update)
CC DE ENV polyprotein (coat polypeptide).
CC GN ENV.
CC OS Simian foamy virus (type 3 / strain LK3) (SFV-3).
CC OC Viruses; Retroid viruses; Retroviridae; Spumavirus.
CC OX NCBI_TaxID=11644;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA MEDLINE=92124734; PubMed=1310187;
CC RA Renne R., Friedl E., Schweizer M., Fleps U., Turek R.,
CC RA Neumann-Haefelin D.;
CC RT "Genomic organization and expression of simian foamy virus type 3

(SFV-3).";
Virology 186:597-608(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74895; AAA47798.1; ALT_INIT.
CC PIR; C40820; VCLJLK.
CC InterPro; IPR005070; Foamy_env.
CC Pfam; PF03408; Foamy_virus_ENV; 1.
CC Coat protein; Transmembrane; Polyprotein; Glycoprotein.
CC TRANSMEM 68 88 I (POTENTIAL).
CC TRANSMEM 955 975 II (POTENTIAL).
CC CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 346 346 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 484 484 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 551 551 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 776 776 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 982 AA; 113313 MW; 721F2F8929D604FF CRC64;

CC Query Match 61.1%; Score 33; DB 1; Length 982;
CC Best Local Similarity 66.7%; Pred. No. 79;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CC QY 1 EEVXPXGMD 9
CC |||||
CC Db 44 EEVPIPTMD 52

CC RESULT 23
CC RPOC_VIBCH
CC ID RPOC_VIBCH STANDARD; PRT; 1401 AA.
CC AC Q9KV29;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
CC DE beta' chain) (RNA polymerase beta' subunit).
CC GN RPOC OR VC0329.
CC OS Vibrio cholerae.
CC OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
CC OX NCBI_TaxID=666;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC SPRAIN=El Tor N16961 / Serotype O1;
CC RX MEDLINE=20406833; PubMed=10952301;
CC RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
CC RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
CC RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
CC RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
CC RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
CC RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
CC RA Fraser C.M.;
CC RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
CC RT cholerae";
CC RL Nature 406:477-483(2000).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA](N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF004121; AAF93502.1; -
 CC HSP; Q9KW6; LHQ.
 CC TIGR: VC0329; -
 CC InterPro: IPR000722; RNA_pol_A.
 CC InterPro: IPR002879; RNA_pol_A2.
 CC Pfam: PF00623; RNA_pol_A; 1.
 CC Pfam: PF01854; RNA_pol_A2; 2.
 CC Transferrase; DNA-directed RNA polymerase; Transcription;
 CC Complete proteome.
 CC KW SEQUENCE 1401 AA; 155021 MW; DFD80F2B514504F CRC64;
 CC
 CC Query Match 61.1%; Score 33; DB 1; Length 1401;
 CC Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 2 EVVPXGMDYS 11
 CC :|||: ||
 CC Db 581 QIVPKGLPYS 590
 CC
 CC RESULT 24
 CC ID ZEP1_HUMAN STANDARD; PRT; 2717 AA.
 CC AC P15822;
 CC DT 01-APR-1990 (Rel. 14, Created)
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)
 CC DT 15-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-
 CC binding protein 1) (HIV-EPI) (Major histocompatibility complex binding
 CC protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
 CC (PRDII-BF1).
 CC GN HIVEP1 OR ZNF40.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP MEDLINE=90169514; PubMed=2106471;
 CC RX Fan C.M., Maniatis T.;
 CC RA "A DNA-binding protein containing two widely separated zinc finger
 CC motifs that recognize the same DNA sequence.";
 CC Genes Dev. 4:29-42(1990).
 CC [2]
 CC RP STRUCTURE BY NMR OF 2113-2142.
 CC RX MEDLINE=91064333; PubMed=2248949;
 CC RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
 CC Gronenborn A.M.;
 CC RA "High-resolution three-dimensional structure of a single zinc finger
 CC from a human enhancer binding protein in solution.";
 CC RL Biochemistry 29:9324-9334(1990).
 CC [3]
 CC RN STRUCTURE BY NMR OF 2087-2142.
 CC RP MEDLINE=92232684; PubMed=1567844;
 CC RX Omichinski J.G., Clore G.M., Roblen M., Sakaguchi K., Appella E.,
 CC Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc
 RT finger from the human enhancer binding protein MBP-1.";
 RL Biochemistry 31:3907-3917(1992).
 CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
 CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
 CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
 CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
 CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS 1
 CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
 CC IN T-CELL ACTIVATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
 CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
 CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
 CC ZINC-FINGER IN-BETWEEN.
 CC -1- SIMILARITY: STRONG, TO HIVEP2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X51435; CAA35798.1; -
 CC PIR: A34203; A34203.
 CC PDB: 3ZNF; 15-JAN-92.
 CC PDB: 4ZNF; 15-JAN-92.
 CC PDB: 1BBO; 31-OCT-93.
 CC TRANSFAC; T00497; -
 CC Genew; HGNC:4920; HIVEP1.
 CC MIM: 194540; -
 CC InterPro: IPR000822; Znf_C2H2.
 CC Pfam: PF00096; zf_C2H2; 5.
 CC PRINTS; PR00048; ZINC-FINGER.
 CC SMART; SM00355; Znf_C2H2; 4.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
 CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 CC Transcription regulation; zinc-finger; Metal-binding; DNA-binding;
 CC Nuclear protein; Repeat; 3D-structure.
 CC KW DOMAIN 406 456 ZINC FINGERS.
 CC FT ZN_FING 406 428 C2H2-TYPE.
 CC FT ZN_FING 434 456 C2H2-TYPE.
 CC FT ZN_FING 434 456 C2H2-TYPE.
 CC FT DOMAIN 806 806 POLY-SER.
 CC FT ZN_FING 958 981 C2HC-TYPE (POTENTIAL).
 CC FT DOMAIN 2087 2139 ZINC FINGERS.
 CC FT ZN_FING 2087 2109 C2H2-TYPE.
 CC FT ZN_FING 2115 2139 C2H2-TYPE.
 CC FT STRAND 2088 2088
 CC FT TURN 2090 2092
 CC FT STRAND 2095 2095
 CC FT HELIX 2099 2108
 CC FT TURN 2109 2109
 CC FT STRAND 2115 2116
 CC FT STRAND 2123 2124
 CC FT HELIX 2127 2135
 CC SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;
 CC
 CC Query Match 61.1%; Score 33; DB 1; Length 2717;
 CC Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 3 VVPXGMDYS 11
 CC |||: ||
 CC Db 2405 VVPAGLYYS 2413
 CC
 CC RESULT 25
 CC ET2A_XENLA STANDARD; PRT; 472 AA.
 CC ID ET2A_XENLA
 CC AC P19102;
 CC DT 01-NOV-1990 (Rel. 16, Created)

Db 49 EQAVPTGLDSYS 60

Search completed: June 10, 2003, 13:40:21
Job time : 4.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 EGVVPMXMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SPTREMBL.21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	74.1	156	3 Q12479	Q12479 saccharomyc
2	38	70.4	363	17 Q30260	Q30260 archaeoglob
3	38	70.4	1063	16 Q8RG86	Q8RG86 fusobacteri
4	36	66.7	341	10 Q22081	Q22081 citrus unsh
5	36	66.7	348	10 Q22096	Q22096 citrus unsh
6	36	66.7	452	10 Q8W568	Q8W568 arabidopsis
7	36	66.7	460	10 Q9C9T7	Q9C9T7 arabidopsis
8	36	66.7	1047	10 P93782	P93782 saccharomyc
9	36	66.7	1083	10 Q9SN30	Q9SN30 arabidopsis
10	36	66.7	1084	10 Q43010	Q43010 oryza sativ
11	36	66.7	1100	10 Q8S064	Q8S064 oryza sativ
12	35	64.8	219	5 Q8G004	Q8G004 eriocheir s
13	35	64.8	253	16 Q8XP88	Q8XP88 clostidium
14	35	64.8	298	2 Q52367	Q52367 rhizobium t
15	35	64.8	425	5 Q9XVK4	Q9XVK4 caenorhabdi
16	35	64.8	433	16 Q9A382	Q9A382 caulobacter

17	35	64.8	440	17 Q9YFI3	Q9YFI3 aeropyrum p
18	35	64.8	511	2 Q52680	Q52680 escherichia
19	35	64.8	517	16 Q8X2L5	Q8X2L5 ralstonia s
20	35	64.8	745	5 Q95P46	Q95P46 carinus ma
21	35	64.8	1031	5 Q9U6A3	Q9U6A3 callinectes
22	35	64.8	1150	5 Q17704	Q17704 caenorhabdi
23	35	64.8	1410	2 Q52673	Q52673 escherichia
24	35	64.8	1420	2 Q52666	Q52666 escherichia
25	35	64.8	1474	17 Q27146	Q27146 methanobact
26	35	64.8	1828	16 Q98K29	Q98K29 rhizobium l
27	35	64.8	2778	5 Q9V9T6	Q9V9T6 drosophila
28	34.5	63.9	748	4 Q8TBJ7	Q8TBJ7 homo sapien
29	34	63.0	143	5 Q9VSX8	Q9VSX8 drosophila
30	34	63.0	154	10 Q9SBB8	Q9SBB8 oryza sativ
31	34	63.0	215	16 Q8R9L5	Q8R9L5 thermoanaer
32	34	63.0	290	16 Q8U7J0	Q8U7J0 agrobacteri
33	34	63.0	296	17 Q9YF88	Q9YF88 aeropyrum p
34	34	63.0	357	17 Q29920	Q29920 archaeoglob
35	34	63.0	366	17 Q29451	Q29451 archaeoglob
36	34	63.0	387	16 Q98FX1	Q98FX1 rhizobium l
37	34	63.0	543	3 Q8TFF4	Q8TFF4 trichoderma
38	34	63.0	558	16 Q8R822	Q8R822 thermoanaer
39	34	63.0	565	16 Q9CIN1	Q9CIN1 lactococcus
40	34	63.0	587	16 Q9JZF8	Q9JZF8 neisseria m
41	34	63.0	906	10 Q9ZSY4	Q9ZSY4 arabidopsis
42	34	63.0	908	10 Q9FJK8	Q9FJK8 arabidopsis
43	34	63.0	908	10 Q8W4J9	Q8W4J9 arabidopsis
44	34	63.0	908	10 Q9ZSV3	Q9ZSV3 arabidopsis
45	34	63.0	909	10 Q9M5A1	Q9M5A1 arabidopsis
46	34	63.0	1062	2 P95422	P95422 pseudomonas
47	34	63.0	1062	16 Q910X8	Q910X8 pseudomonas
48	34	63.0	3472	1 Q74056	Q74056 cenarchaeum
49	33	61.1	78	6 Q9XST4	Q9XST4 canis famil
50	33	61.1	97	17 Q97VR9	Q97VR9 sulfolobus
51	33	61.1	128	17 Q97US8	Q97US8 sulfolobus
52	33	61.1	172	13 Q02528	Q02528 oryzias lat
53	33	61.1	175	10 Q8VY88	Q8VY88 arabidopsis
54	33	61.1	183	2 Q9S1I0	Q9S1I0 sweet potat
55	33	61.1	184	17 Q97XJ2	Q97XJ2 sulfolobus
56	33	61.1	193	5 Q8SWL4	Q8SWL4 encephalito
57	33	61.1	209	13 Q9DEL6	Q9DEL6 brachydanio
58	33	61.1	210	10 Q6S890	Q6S890 cyclorella
59	33	61.1	217	4 Q00404	Q00404 homo sapien
60	33	61.1	225	10 Q40129	Q40129 lycopersico
61	33	61.1	230	17 Q9P9L6	Q9P9L6 pyrobaculum
62	33	61.1	247	16 Q92U66	Q92U66 rhizobium m
63	33	61.1	257	10 Q9C6J0	Q9C6J0 arabidopsis
64	33	61.1	262	17 Q97YB8	Q97YB8 sulfolobus
65	33	61.1	267	17 Q97Y57	Q97Y57 sulfolobus
66	33	61.1	299	4 Q90EE9	Q90EE9 homo sapien
67	33	61.1	299	17 Q97TV4	Q97TV4 sulfolobus
68	33	61.1	299	17 Q9UWN4	Q9UWN4 sulfolobus
69	33	61.1	307	17 Q9HP60	Q9HP60 halobacteri
70	33	61.1	396	17 Q8TIT9	Q8TIT9 methanosarc
71	33	61.1	441	17 Q9HIW6	Q9HIW6 thermoplasma
72	33	61.1	556	4 Q43733	Q43733 homo sapien
73	33	61.1	577	15 Q9QBR8	Q9QBR8 simian foam
74	33	61.1	577	15 Q9QBR7	Q9QBR7 simian foam
75	33	61.1	577	15 Q9QBR6	Q9QBR6 simian foam

ALIGNMENTS

RESULT 1
Q12479 PRELIMINARY; PRT; 156 AA.
ID Q12479
AC Q12479
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Q9XVK4 caenorhabdi
GN YOR013W.

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA De haan M., Grivell L.A., Maarse A.C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN=FY1679;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA STRAIN=FY1679;
 RL MEDLINE=94019318; PubMed=8413243;
 RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
 RA Sherman F.;
 RL "CYC2 encodes a factor involved in mitochondrial import of yeast
 RL cytochrome c.";
 RL Mol. Cell. Biol. 13:6442-6451(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA STRAIN=FY1679;
 RL MEDLINE=94169519; PubMed=7764548;
 RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
 RL "Molecular cloning of a gene, DHS1, which complements a drug-
 RL hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
 RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
 DR EMBL; Z74920; CAA99201.1; -;
 DR EMBL; X87331; CAA60762.1; -;
 DR SGD; S0005539; YOR013W.
 SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 74.1%; Score 40; DB 3; Length 156;
 Best Local Similarity 77.8%; Pred. No. 1.7;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
 Db 50 EVWPLGMDY 58
 ||| |||||

RESULT 2
 O30260
 ID O30260 PRELIMINARY; PRT; 363 AA.
 AC O30260;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein AF2411.
 GN AF2411.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kane B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RL reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91255.1; -;
 DR TIGR; AF2411; -;
 DR InterPro; IPR002103; Bac_luciferase.
 DR Pfam; PF00296; bac_luciferase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 363 AA; 41736 MW; 0E976EAE788F4803 CRC64;

Query Match 70.4%; Score 38; DB 17; Length 363;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 Db 120 ENIVPYGIDFS 130
 ||| |||||

RESULT 3
 Q8RG86
 ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
 AC Q8RG86;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
 GN FN0422.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010554; AAL94625.1; -;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDYS 11
 Db 195 EIVPGLNYS 204
 ||| |||||

RESULT 4
 Q22081
 ID Q22081 PRELIMINARY; PRT; 341 AA.
 AC Q22081;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Sucrose-phosphate synthase (Fragment).
 GN CITSPS2.
 OS Citrus unshiu (Satsuma orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=55188;

RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
phosphate synthase isoforms from a citrus fruit (Citrus unshiu
Marc.).";
RL Mol. Gen. Genet. 252:346-351(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR EMBL; AB006319; BAA23215.1; -.
FT NON_TER 1 341
FT SEQUENCE 341 AA; 38136 MW; 61417A69C4560777 CRC64;
SQ
Query Match 66.7%; Score 36; DB 10; Length 341;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMDYS 11
DB 228 VVPGMDFS 236
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
phosphate synthase isoforms from a citrus fruit (Citrus unshiu
Marc.).";
RL Mol. Gen. Genet. 252:346-351(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR EMBL; AB006660; BAA22071.1; -.
FT NON_TER 1 348
FT SEQUENCE 348 AA; 38556 MW; EEIC21EBA6FF5C5E CRC64;
SQ
Query Match 66.7%; Score 36; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMDYS 11
DB 234 VVPGMDFS 242

RESULT 6
Q8W568 PRELIMINARY; PRT; 452 AA.
ID Q8W568;
AC Q8W568;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Atlg73750/F25P22.17
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF419606; AAL31938.1; -.
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00708; PRO-ENDOPEP_SER; UNKNOWN_1.
SQ SEQUENCE 452 AA; 49682 MW; A159955B21742C4A CRC64;
Query Match 66.7%; Score 36; DB 10; Length 452;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
DB 210 EEDVPSAMDY 219
[1]
RESULT 7
Q9C9T7 PRELIMINARY; PRT; 460 AA.
ID Q9C9T7;
AC Q9C9T7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 50.6 kDa protein.
GN F25P22.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

```

RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC012679; AAG52073.1; -.
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 50564 MW; E94B27B5C4B249EC CRC64;

Query Match 66.7%; Score 36; DB 10; Length 460;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
Db 218 EEDVPSAMDY 227
II III

RESULT 8
P93782 PRELIMINARY; PRT; 1047 AA.
AC P93782;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (Fragment).
GN SOSPS1.
OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]
RP TISSUE=LEAF;
RC Sugiharto B., Sakakibara H., Sugiyama T.;
RT "Differential Expression of Two Genes for Sucrose-Phosphate Synthase
RT in Sugarcane: Molecular Cloning of the cDNAs and Comparative Analysis
RT of Gene Expression.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB001337; BAA19241.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 1047 AA; 116379 MW; D0EDB34961E1D83D CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1047;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 414 VIPPGMDFS 422
I: I I I I I

RESULT 9
Q9SN30 PRELIMINARY; PRT; 1083 AA.
ID Q9SN30
AC Q9SN30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase-like protein (EC 2.4.1.14).
CN F28W1.40 OR AT4G10120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```

```

[1] SEQUENCE FROM N.A.
RN RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2] SEQUENCE FROM N.A.
RN RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[3] SEQUENCE FROM N.A.
RN RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049487; CAB39764.1; -.
DR EMBL; AL161516; CAB78135.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1083 AA; 122688 MW; EFC126FCA2137BB1 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 483 VIPPGMDFS 491
I: I I I I I

RESULT 10
Q43010 PRELIMINARY; PRT; 1084 AA.
ID Q43010;
AC Q43010;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).
GN SP51.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. JAPONICA;
RA Sakamoto M., Satozawa T., Kishimoto N., Higo K., Shimada H.,
RA Fujimura T.;
RT "Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS)
RT gene that is specifically expressed in the source organ.";
RL Plant Sci. 112:207-217(1995).
CC -!- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) -> PYRUVATE + NADH.
CC -!- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
DR EMBL; D45890; BAA08304.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR001557; L_LDH.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR PROSITE; PS00064; L_LDH; 1.
KW Glycolysis; NAD: Oxidoreductase.
SQ SEQUENCE 1084 AA; 119417 MW; 4D0D79AFEDD2F92B CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 453 VIPPGMDFS 461
I: I I I I I

RESULT 11
Q8S064

```

```

ID Q8S064 PRELIMINARY; PRT; 1100 AA.
AC Q8S064;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE putative sucrose-phosphate synthase.
GN P0678F11.14.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriocaulaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0678F11." to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003437; BAB86107.1; -.
SQ SEQUENCE 1100 AA; 121170 MW; 084F4604BA389CAD CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1100;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 469 VIPPGMDFS 477

RESULT 12
Q9GQ04 PRELIMINARY; PRT; 219 AA.
AC Q9GQ04;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Na+/K+/2Cl-cotransporter (Fragment).
OS Eriocheir sinensis (Chinese mitten crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Grapsoidae; Varunidae; Eriocheir.
OX NCBI_TaxID=95602;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GILL;
RA Weihrach D., Towle D.W.;
RT "Na+/H+-exchanger and Na+/K+/2Cl- cotransporter are expressed in
gills of the euryhaline Chinese crab Eriocheir sinensis.";
RL Comp. Biochem. Physiol. 126:SI58-SI58(2000).
DR EMBL; AF301160; AAG39938.1; -.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 24159 MW; 599442DA26F0D3F1 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 219;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VPXGMDYS 11
Db 107 VPOGLDYS 114

RESULT 13
Q8XPA8 PRELIMINARY; PRT; 253 AA.
AC Q8XPA8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein CPE0057.

```

```

GN CPE0057.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003185; BAB79763.1; -.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR000594; THF_domain.
DR Pfam; PF00899; Thif; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 253 AA; 27714 MW; 814DF79D3E0D7486 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 253;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
Db 108 EEIPDDVDY 117

RESULT 14
Q52367 PRELIMINARY; PRT; 298 AA.
AC Q52367;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Aryl-alcohol dehydrogenase homolog (Fragment).
GN XYL81.
OS Rhizobium tropici.
OG Plasmid pRCFN299a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=398;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN299;
RA Rosenblueth M., Hynes M.F., Martinez-Romero E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
FAMILY.
DR EMBL; AF036920; AAC04779.1; -.
DR HSSP; P07846; LSDG.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Plasmid; Zinc.
FT NON_TER 298
FT NON_TER 298
SQ SEQUENCE 298 AA; 31092 MW; 49B2F8117C33AE87 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 298;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGMDYS 11
Db 250 EIIPEGADFS 259

RESULT 15

```

```

Query Match          64.8%; Score 35; DB 16; Length 433;
Best Local Similarity 54.5%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMGMYD 11
      ||:|:|
Db      266 EVLPPGFDYS 276

RESULT 17
Q9YFT3
ID Q9YFI3 PRELIMINARY; PRT; 440 AA.
AC Q9YFI3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 440AA long hypothetical alkaline protease.
GN APE0263.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000058; BAA79178.1; -
DR HSP; P00782; 2SPT.
DR InterPro: IPR00209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 440 AA; 44680 MW; E43E6E2174B6F07E CRC64;

Query Match          64.8%; Score 35; DB 17; Length 440;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 EVVPXGMGYD 10
      ||:|:|
Db      120 EVLPWGVYD 128

RESULT 18
OS2680 PRELIMINARY; PRT; 511 AA.
ID OS2680; Q47289;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
GN RHSG.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR-50;
RA Wang Y.-D., Zhao S., Hill C.W.;

```

```
RT "Rhs elements comprise three subfamilies.";
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF044506; AAC622388.1; -.
FT NON_TER 511
SQ SEQUENCE 511 AA; 55877 MW; C85402569450DBFF CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 2; Length 511;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVVPXGMDY 10
Db 381 EQVNPGLDY 390

RESULT 19
Q8XZL5 PRELIMINARY; PRT; 517 AA.
AC Q8XZL5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Probable SUBSTATE-binding periplasmic (PBP) ABC transporter
DE protein.
GN RSC1380 OR RS04663.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Laviue M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646064; CAD15082.1; -.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
KW Complete proteome.
SQ SEQUENCE 517 AA; 57349 MW; 861EE836F645F2DA CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 16; Length 517;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 316 VVPQGVDA 324

RESULT 20
Q95P46 PRELIMINARY; PRT; 745 AA.
AC Q95P46;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Na+/K+/2Cl-cotransporter (Fragment).
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GILL;
RA Wehrauch D., Towle D.W.;

"Sequence and expression analysis of the Na+/K+/2Cl- cotransporter in
the euryhaline crabs Callinectes sapidus and Carcinus maenas.";
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035548; AAK62044.1; -.
DR InterPro: IPR002293; AA/rei_prmeasel.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
FT NON_TER 1
FT NON_TER 745
SQ SEQUENCE 745 AA; 81249 MW; 08AE2D3E1F70C7BE CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 5; Length 745;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMDYS 11
Db 625 VPOGLDYS 632

RESULT 21
Q9U6A3 PRELIMINARY; PRT; 1031 AA.
AC Q9U6A3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Na+/K+/2Cl-cotransporter.
DE NKCC.
GN Callinectes sapidus (Blue crab).
OS Callinectes sapidus (Blue crab); Crustacea; Malacostraca;
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Callinectes.
OX NCBI_TaxID=6763;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GILL;
RA Towle D.W.;
RA "Cloning and sequencing a Na+/K+/2Cl- cotransporter from gills of the
euryhaline blue crab Callinectes sapidus.";
RL Am. Zoologist 38:114A-114A(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=GILL;
RA Towle D.W., Wehrauch D.;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF190129; AAF05702.1; -.
DR InterPro: IPR002293; AA/rei_prmeasel.
DR InterPro: IPR004842; KCl_cotransport.
DR InterPro: IPR002443; NaKCl_transporter.
DR PRINTS: PR01207; NAKCLTRNSPRT.
DR TIGRfams: TIGR00930; 2a30; 1.
SQ SEQUENCE 1031 AA; 113699 MW; F67773021AB05D71 CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 5; Length 1031;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMDYS 11
Db 737 VPOGLDYS 744

RESULT 22
Q17704 PRELIMINARY; PRT; 1150 AA.
AC Q17704; 017976;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE C53A5.2 protein.
GN C53A5.2.
OS Caenorhabditis elegans.
```

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Mortimore B.;
 RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Bonfield J., Alnsough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Letellier P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin A., Roopra A., Saunders D., Shownkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans".
 RL Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 281486; CAB03994.1; -;
 DR EMBL; 278015; CAB03994.1; JOINED.
 DR EMBL; 278015; CAB01437.1; -;
 DR EMBL; 281486; CAB01437.1; JOINED.
 DR InterPro; IPR001201; PAP_25A.Core.
 DR InterPro; IPR000051; SAM_bind.
 DR InterPro; IPR003402; Unk_Met10.
 DR Pfam; PF02475; Met_10; I.
 SQ SEQUENCE 1150 AA; 133001 MW; BAI30A251C1A12F8 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 1150;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 VVPXGMDYS 11
 DB 562 VLPVGIDYS 570
 | : | : | : |

RESULT 23
 O52673
 ID 052673 PRELIMINARY; PRT; 1410 AA.
 AC 052673;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Core protein.
 DE Putative membrane protein.
 GN MTH1074.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC11;
 RX MEDLINE=98361897; PubMed=9696756;
 RA Wang Y.D., Zhao S., Hill C.W.;
 RT "Rhs elements comprise three subfamilies which diverged prior to
 RT acquisition by Escherichia coli.";
 RL J. Bacteriol. 180:4102-4110(1998).
 DR EMBL; AF044503; AAC32478.1; -;
 DR InterPro; IPR001826; RHS.
 DR Pfam; PF03527; RHS; 1.
 DR PRINTS; PR00394; RHSPROTEIN.
 SQ SEQUENCE 1410 AA; 158739 MW; 39ECB9F8C97FCB15 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 1410;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EEVVPXGMDY 10
 DB 382 EQVNPEGIDY 391
 | : | : | : |

RESULT 24
 O52666
 ID 052666 PRELIMINARY; PRT; 1420 AA.
 AC 052666;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Core protein.
 DE Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC45;
 RX MEDLINE=98361897; PubMed=9696756;
 RA Wang Y.D., Zhao S., Hill C.W.;
 RT "Rhs elements comprise three subfamilies which diverged prior to
 RT acquisition by Escherichia coli.";
 RL J. Bacteriol. 180:4102-4110(1998).
 DR EMBL; AF044501; AAC32471.1; -;
 DR InterPro; IPR001826; RHS.
 DR InterPro; IPR00130; Zn_MTpeptdse.
 DR Pfam; PF03527; RHS; 1.
 DR PRINTS; PR00394; RHSPROTEIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1420 AA; 159215 MW; E7851E4D48740621 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 1420;
 Best Local Similarity 60.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EEVVPXGMDY 10
 DB 381 EQVNPEGIDY 390
 | : | : | : |

RESULT 25
 O27146
 ID 027146 PRELIMINARY; PRT; 1474 AA.
 AC 027146;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative membrane protein.
 GN MTH1074.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiawani N., Caruso A., Bush D., Safer H., Pietrovski S., Church G.M.,
 RA McDougall S., Shimer G., Goyal A., Rice P., Noelling J., Reeve J.N.;
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AB000878; AAB85563.1; -;

DR InterPro; IPR001434; DUF11.
DR Pfam; PF01345; DUF11; 9.
KW Complete proteome.
SQ SEQUENCE 1474 AA; 153713 MW; 11D4D27BD86255F0 CRC64;

Query Match 64.8%; Score 35; DB 17; Length 1474;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
: : : : :
Db 1238 DVLPAGLDY 1246

Search completed: June 10, 2003, 13:46:34
Job time : 26.7857 secs

Best Local Similarity 70.0%; Pred. No. 2.6e+02; DB 9: Length 736;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVYGM DY 340

RESULT 22

US-10-173-706-420
; Sequence 420, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-420

Query Match 61.1%; Score 33; DB 9: Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVYGM DY 340

RESULT 23

US-10-175-738-420
; Sequence 420, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-175-738-420

Query Match 61.1%; Score 33; DB 9: Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVYGM DY 340

RESULT 24

US-10-175-752-420
; Sequence 420, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-420

Query Match 61.1%; Score 33; DB 9: Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVYGM DY 340

RESULT 25

US-10-176-482-420
; Sequence 420, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 420

; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-420

Query Match 61.1%; Score 33; DB 9; Length 736;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| | | | |
Db 331 EPVVVYGM DY 340

Search completed: June 10, 2003, 14:35:43
Job time : 15.0714 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds
(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-12
Perfect score: 54
Sequence: 1 EEVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	40	74.1	156	S54619	hypothetical prote
2	38	70.4	363	D69551	conserved hypothet
3	36	66.7	102	A42452	VI protein - tobac
4	36	66.7	341	S72649	sucrose-phosphate
5	36	66.7	348	S72650	sucrose-phosphate
6	36	66.7	460	G96764	unknown protein F2
7	36	66.7	1049	JC4783	sucrose-phosphate
8	36	66.7	1068	JQ1329	sucrose-phosphate
9	36	66.7	1081	T09837	sucrose-phosphate
10	36	66.7	1083	T04062	sucrose-phosphate
11	36	66.7	1084	T04103	sucrose-phosphate
12	35	64.8	425	T24111	hypothetical prote
13	35	64.8	433	H87660	peptidoglycan-bind
14	35	64.8	440	H27284	probable alkaline
15	35	64.8	1150	T20173	hypothetical prote
16	35	64.8	1474	F69009	probable membrane
17	35	64.8	2747	B49132	fat facets (faf) s
18	34	63.0	99	S00210	plastoeyanin b - L
19	34	63.0	155	S38255	plastoeyanin b pre
20	34	63.0	168	S58208	plastoeyanin b
21	34	63.0	290	D98182	6-O-methylguanine-D
22	34	63.0	290	AG3104	6-O-methylguanine-
23	34	63.0	296	F72745	hypothetical prote
24	34	63.0	357	G69290	probable hexosyltr
25	34	63.0	366	T26350	L-lactate dehydrog
26	34	63.0	565	G68665	ABC transporter AT
27	34	63.0	566	A70164	phenylalanine-tRNA
28	34	63.0	587	F81138	succinate dehydrog
29	34	63.0	906	T48898	disease resistance

ALIGNMENTS

RESULT 1

S54619
hypothetical protein YOR013w - yeast (Saccharomycetes cerevisiae)
N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3
C:Species: Saccharomycetes cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
C:Accession: S54619; S66879
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.
A:Reference number: S54617
A:Accession: S54619
A:Molecule type: DNA
A:Residues: 1-156 <DEH>
A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.
A:Reference number: S54617
A:Accession: S54619
A:Molecule type: DNA
A:Residues: 1-156 <DEH>
A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:
A:Experimental source: strain S288C

disease resistance
RND multidrug effl
hypothetical prote
hypothetical 367K
partial transposas
hypothetical prote
fibroblast growth
hypothetical prote
hypothetical prote
conserved hypothet
unknown protein [1
transposase ISC105
transposase ISC105
pantoate-beta-alan
transposase ISC105
transposase ISC105
hypothetical prote
acyl-CoA dehydroge
ferrisiderophore r
zinc finger protei
fibroblast growth
heparin-binding gr
fibroblast growth
protein-tyrosine k
penicillin-binding
probable sulfate p
penicillin-binding
sulfate permease -
env polyprotein -
hypothetical prote
DNA-directed RNA p
DNA-binding protei
cyclic beta 1-2 gl
cellobiose-phospho
transcription fact
plastoeyanin - car
thiol peroxidases
thiol peroxidases
transcription fact
fimbrial chain [1m
L-ribulose-phospha
L-ribulose-5-phosp
L-ribulose-5-phosp
L-ribulose-5-phosp

30 34 63.0 908 2 T48899
31 34 63.0 1062 2 F83335
32 34 63.0 1062 2 T30830
33 34 63.0 3472 2 T31308
34 33 61.1 97 2 A94427
35 33 61.1 128 2 A90471
36 33 61.1 172 2 S27021
37 33 61.1 184 2 E90335
38 33 61.1 225 2 S57810
39 33 61.1 247 2 A66001
40 33 61.1 257 2 A96546
41 33 61.1 262 2 F90298
42 33 61.1 267 2 C90307
43 33 61.1 276 2 C64417
44 33 61.1 283 2 G83055
45 33 61.1 299 2 E90487
46 33 61.1 299 2 H90352
47 33 61.1 307 2 F84330
48 33 61.1 394 2 F82491
49 33 61.1 421 1 DERTCM
50 33 61.1 670 2 S22293
51 33 61.1 797 2 S38579
52 33 61.1 800 1 TVH02F
53 33 61.1 800 2 A48991
54 33 61.1 801 2 I55363
55 33 61.1 806 2 A35963
56 33 61.1 840 2 AG0526
57 33 61.1 840 2 T39116
58 33 61.1 846 2 S57580
59 33 61.1 877 2 T40413
60 33 61.1 982 1 VCLJLK
61 33 61.1 1064 2 F86182
62 33 61.1 1401 2 G82336
63 33 61.1 2717 2 A34203
64 33 61.1 2831 2 T31419
65 33 61.1 2867 2 AG3481
66 32.5 60.2 472 1 B53236
67 32 59.3 97 2 JW0011
68 32 59.3 165 2 AG1272
69 32 59.3 165 2 AH1635
70 32 59.3 175 2 S36749
71 32 59.3 180 2 AG0504
72 32 59.3 231 1 ISECP4
73 32 59.3 231 2 A90637
74 32 59.3 231 2 A85488
75 32 59.3 231 2 AB0515

C:Genetics:
 A:Cross-references: SGD:S0005539
 A:Map position: 15R
 C:Superfamily: hypothetical protein YOR013W

Query Match 74.1%; Score 40; DB 2; Length 156;
 Best Local Similarity 77.8%; Pred. No. 1;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
 ||:| ||||
 Db 50 EVMPGLGMDY 58

RESULT 2

D69551
 conserved hypothetical protein AF2411 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: D69551
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: D69551
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-363 <KLE>
 A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AAB91255.1; PID:g265068

Query Match 70.4%; Score 38; DB 2; Length 363;
 Best Local Similarity 54.5%; Pred. No. 6.8;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 | : || | : | : |
 Db 120 ENIVPXGIDFS 130

RESULT 3

A42452
 V1 protein - tobacco yellow dwarf virus (strain Australia)
 C:Species: tobacco yellow dwarf virus
 C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
 C:Accession: A42452
 R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
 Virology 187, 633-642, 1992
 A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yell
 A:Reference number: A42452; MUID:92188538; PMID:1546458
 A:Accession: A42452
 A:Molecule type: DNA
 A:Residues: 1-102 <MOR>
 A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 66.7%; Score 36; DB 2; Length 102;
 Best Local Similarity 60.0%; Pred. No. 4.2;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDYS 11
 : || | | : | : |
 Db 7 QVVPGGINYS 16

RESULT 4

S72649
 sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Citrus unshiu (fragment)
 C:Species: Citrus unshiu
 C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
 C:Accession: S72649

R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
 Mol. Gen. Genet. 252, 346-351, 1996
 A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate sy
 A:Reference number: S72648; MUID:96439842; PMID:8842155

A:Accession: S72649
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-341 <KOM>

A:Cross-references: EMBL:AB006319; NID:g2588891; PIDN:BAA23215.1; PID:g2588892
 A:Experimental source: fruit, cv. Miyagawa-Wase
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997
 C:Genetics:
 A:Gene: SPS2

C:Function:

A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fruc
 A:Pathway: sucrose biosynthesis
 C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homolo
 C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
 F:1-341/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 341;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 | : | | | : |
 Db 228 VIPPGMDFS 236

RESULT 5

S72650
 sucrose-phosphate synthase (EC 2.4.1.14) isoform 3 - Citrus unshiu (fragment)
 C:Species: Citrus unshiu
 C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
 C:Accession: S72650
 R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
 Mol. Gen. Genet. 252, 346-351, 1996
 A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate sy
 A:Reference number: S72648; MUID:96439842; PMID:8842155
 A:Accession: S72650
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-348 <KOM>

A:Cross-references: EMBL:AB006660; NID:g2351059; PIDN:BAA22071.1; PID:g2351060
 A:Experimental source: fruit, cv. Miyagawa-Wase
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997
 C:Genetics:
 A:Gene: SPS3

C:Function:

A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fruc
 A:Pathway: sucrose biosynthesis
 C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homolo
 C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
 F:1-348/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 348;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 | : | | | : |
 Db 234 VIPPGMDFS 242

RESULT 6

G96764
 unknown protein F25P22.17 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: G96764
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizlar, L.

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khavkin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96764
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-460 <STO>
A:Cross-references: GB:AE0051173; NID:g6692750; PIDN:AAF24856.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25P22.17
A:Map position: 1

Query Match 66.7%; Score 36; DB 2; Length 460;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
|| || |||
Db 218 EEDVPSAMDY 227

RESULT 7
JC4783
sucrose-phosphate synthase (EC 2.4.1.14) - rice
C:Species: Oryza sativa (rice)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 18-Jun-1999
C:Accession: JC4783
R:Valdez-Alarcon, J.J.; Ferrando, M.; Salerno, G.; Jimenez-Moralla, B.; Herrera-Estrella
Gene 170, 217-222, 1996
A:Title: Characterization of a rice sucrose-phosphate synthase-encoding gene.
A:Reference number: JC4783; MUID:96235138; PMID:8666248
A:Accession: JC4783
A:Molecule type: mRNA
A:Residues: 1-1049 <VAL>
A:Cross-references: GB:U33175; NID:g1449931; PIDN:ACA93379.1; PID:g988270
A:Note: UDPglucose-fructose-phosphate glucosyltransferase; Sucrosephosphate-UDPglucosyl
C:Comment: This enzyme catalyzes the formation of sucrose-phosphate form UDP-glucose and
C:Genetics:
A:Gene: Sps1
A:Introns: 24/1; 103/3; 183/3; 205/3; 435/3; 475/3; 519/3; 578/3; 596/3; 617/3; 931/3; 9
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:180-663/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1049;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|| || |||
Db 436 VIPPGMDFS 444

RESULT 8
JQ1329
sucrose-phosphate synthase (EC 2.4.1.14) - maize
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JQ1329; PQ0260
R:Worrell, A.C.; Bruneau, J.M.; Summerfelt, K.; Boersig, M.; Voelker, T.A.
Plant Cell 3, 1121-1130, 1991
A:Title: Expression of a maize sucrose phosphate synthase in tomato alters leaf carbohyd
A:Reference number: JQ1329; MUID:92338837; PMID:1840396
A:Accession: JQ1329
A:Molecule type: mRNA

A:Residues: 1-1068 <WOR>
A:Cross-references: GB:M97550; NID:g168625; PIDN:AAA33513.1; PID:g168626
A:Accession: PQ0260
A:Molecule type: protein
A:Residues: 71-74; 206-212; 471-481; 872-892 <WOR1>
C:Comment: This enzyme transfers the glucosyl group from UDPglucose to fructose-6-p
C:Comment: This enzyme is involved in the regulation of carbon partitioning in the
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D
A:Pathway: sucrose biosynthesis
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:178-666/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 1; Length 1068;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|| || |||
Db 435 VIPPGMDFS 443

RESULT 9
T09837
sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Craterostigma plantagineum
C:Species: Craterostigma plantagineum
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09837
R:Ingram, J.; Chandler, J.W.; Gallagher, L.; Salamini, F.; Bartels, D.
Plant Physiol. 115, 113-121, 1997
A:Title: Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to
A:Reference number: Z18874; MUID:97451773; PMID:9306694
A:Accession: T09837
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1081 <ING>
A:Cross-references: EMBL:Y11795; NID:g2190349; PIDN:CAA72491.1; PID:g2190350
A:Experimental source: ABA-treated callus
C:Genetics:
A:Gene: sps2
C:Function:
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:176-674/Domain: sucrose/sucrose-phosphate synthase homology <SSS>

Query Match 66.7%; Score 36; DB 2; Length 1081;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|| || |||
Db 445 VIPPGMDFS 453

RESULT 10
T04062
sucrose-phosphate synthase homolog F28M11.40 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 16-Jul-1999
C:Accession: T04062
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15184
A:Accession: T04062
A:Molecule type: DNA
A:Residues: 1-1083 <BEV>
A:Cross-references: EMBL:AL049487
A:Experimental source: cultivar Columbia; BAC clone F28M11
C:Genetics:
A:Map position: 4

A;Status: preliminary; transl

A:Molecule type: DNA
 A:Residues: 1-1150 <WIL>
 A:Cross-references: EMBL:281486; PIDN:CAB03994.1; GSPDB:GN00023; CESP:C53A5.2
 A:Experimental source: clone C53A5
 R:Matthews, L.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: Z19808
 A:Accession: T23857
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1150 <WIL>
 A:Cross-references: EMBL:278015; PIDN:CAB01437.1; GSPDB:GN00023; CESP:C53A5.2
 A:Experimental source: clone R02D5
 C:Genetics:
 A:Gene: CESP:C53A5.2
 A:Map position: 5
 A:Introns: 33/3; 63/3; 132/3; 169/3; 221/3; 299/3; 379/2; 423/2; 438/2; 471/1; 513/2; 65

Query Match 64.8%; Score 35; DB 2; Length 1150;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 I: I: I: I: I:
 Db 562 VLPVGIDYS 570

RESULT 16
 F69009
 probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta H
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: F69009
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: F69009
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1474 <MTH>
 A:Cross-references: GB:AE000878; GB:AE000666; NID:32622171; PIDN:AAB85563.1; PID:3262217
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1074
 C:Keywords: duplication

Query Match 64.8%; Score 35; DB 2; Length 1474;
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
 I: I: I: I: I:
 Db 1238 DVLPAGLDY 1246

RESULT 17
 B49132
 fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
 C:Accession: B49132; A49132
 R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
 Development 116, 985-1000, 1992
 A:Title: The fat facets gene is required for Drosophila eye and embryo development.
 A:Reference number: A49132; MUID:93202020; PMID:1295747
 A:Contents: isogenic st
 A:Accession: B49132
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-2747 <FIS>

A:Cross-references: GB:L04959; NID:g157411; PIDN:AAF01345.1; PID:g6013474
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
 A:Accession: A49132
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-2704, 'VT', 2707, 'ANNV' <FI2>
 A:Cross-references: GB:L04958; NID:g157410; PIDN:AAF01346.1; PID:g6013475
 A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
 C:Keywords: alternative splicing

Query Match 64.8%; Score 35; DB 2; Length 2747;
 Best Local Similarity 54.5%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: I: I: I: I:
 Db 1394 EVIVPDGODFS 1404

RESULT 18
 S00210
 plastocyanin b - Lombardy poplar
 C:Species: Populus nigra var. italica (Lombardy poplar)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 20-Apr-2000
 C:Accession: S00210
 R:Dimitrov, M.I.; Egorov, C.A.; Donchev, A.A.; Atanasov, B.P.
 FEBS Lett. 226, 17-22, 1987
 A:Title: Complete amino acid sequence of poplar plastocyanin b.
 A:Reference number: S00210
 A:Accession: S00210
 A:Molecule type: protein
 A:Residues: 1-99 <DIM>
 C:Superfamily: plastocyanin
 C:Keywords: chloroplast; copper; electron transfer; metalloprotein
 F:37,84,87,92/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 63.0%; Score 34; DB 2; Length 99;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: I: I: I: I:
 Db 43 EDAPSPGVDVS 53

RESULT 19
 S38255
 plastocyanin precursor - barley
 C:Species: Hordeum vulgare (barley)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jun-1999
 C:Accession: S38255; S00206
 R:Nielsen, P.S.; Gausing, K.
 Eur. J. Biochem. 217, 97-104, 1993
 A:Title: In vitro binding of nuclear proteins to the barley plastocyanin gene prom
 A:Reference number: S38255; MUID:94039081; PMID:8223592
 A:Accession: S38255
 A:Molecule type: DNA
 A:Residues: 1-155 <NIE1>
 A:Cross-references: EMBL:228347; NID:g431919; PIDN:CAA82201.1; PID:g431920
 A:Experimental source: strain NK 1558
 R:Nielsen, P.S.; Gausing, K.
 FEBS Lett. 225, 159-162, 1987
 A:Title: The precursor of barley plastocyanin: sequence of cDNA clones and gene exp
 A:Reference number: S00206
 A:Accession: S00206
 A:Molecule type: mRNA
 A:Residues: 1-119, 'T', 121-155 <NIE2>
 A:Cross-references: EMBL:Y00704; NID:g22704
 A:Note: not compared to nucleotide translation
 C:Genetics:
 A:Genome: nuclear

C;Function: mediates the transfer of electrons from cytochrome b6/f to photosystem I
C;Superfamily: plastocyanin
C;Keywords: chloroplast; copper; electron transfer; metalloprotein
F:1-58/Domain: transit peptide (chloroplast) #status: predicted <NP>
F:59-155/Product: plastocyanin #status: predicted <MAR>
F:95,140,143,148/Binding site: copper (His, Cys, His, Met) (type 1) #status: predicted

Query Match 63.0%; Score 34; DB 2; Length 155;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
| : | | | : | |
Db 101 EDAVPSGVDYS 111

RESULT 20
S58208
Plastocyanin b precursor - black poplar
C;Species: Populus nigra (black poplar)
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-May-2000
C;Accession: S58208
R;Reichert, J.; Jenzelewski, V.; Haehnel, W.
submitted to the EMBL Data Library, July 1995

A;Description: Kinetic studies of recombinant poplar plastocyanins.
A;Reference number: S58208
A;Accession: S58208
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-168 <PEL>
A;Cross-references: EMBL:Z50186; NID:929814; PIDN:CAA90565.1; PID:929815
C;Superfamily: plastocyanin
C;Keywords: copper; electron transfer; metalloprotein
F:106,153,156,161/Binding site: copper (His, Cys, His, Met) (type 1) #status: predicted

Query Match 63.0%; Score 34; DB 2; Length 168;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
| : | | | : | |
Db 112 EDAVPSGVDYS 122

RESULT 21
D98182
O6-methylguanine-DNA methyltransferase PA2118 [imported] - Agrobacterium tumefaciens (st
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: D98182
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: D98182
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88982.1; PID:915158766; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_818
A;Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
| : | | | : | |
Db 9 EDITPTIGSDY 18

RESULT 22

AG3104

6-O-methylguanine-DNA methyltransferase [imported] - Agrobacterium tumefaciens (str
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AG3104
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; W
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; Mc
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58
A;Reference number: AB2577; PMID:11743193
A;Accession: AG3104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45253.1; PID:917742937; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: ada
A;Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10

| : | | | : | |

Db 9 EDITPTIGSDY 18

RESULT 23

F72745

hypothetical protein APE0493 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C;Accession: F72745
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; T
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer
A;Reference number: A72450; PMID:99310339; PMID:10382966
A;Accession: F72745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <KAW>
A;Cross-references: DDBJ:AP000059; NID:95103911; PIDN:BAA79458.1; PID:91043244; PID:
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0493
C;Superfamily: Aeropyrum pernix hypothetical protein APE0493

Query Match 63.0%; Score 34; DB 2; Length 296;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMDYS 11

| : | | | : | |

Db 3 ETLPGGLDYT 12

RESULT 24

G69290

probable hexosyltransferase (EC 2.4.1.-) AF0327 [similarity] - Archaeoglobus fulgidu
C;Species: Archaeoglobus fulgidus
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: G69290
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; DO
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69290
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-357 <KLE>
 A:Cross-references: GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AAB90909.1; PID:g265031
 C:Superfamily: probable hexosyltransferase ycxN
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 63.0%; Score 34; DB 1; Length 357;
 Best Local Similarity 55.6%; Pred. No. 44;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
 DB 170 EVIPNGIDF 178

RESULT 25

G69350
 L-lactate dehydrogenase, cytochrome-type (lldD) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 17-Mar-2000
 C:Accession: G69350
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69350
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-366 <KLE>
 A:Cross-references: GB:AE001049; GB:AE000782; NID:g2689372; PIDN:AAB90435.1; PID:g264980
 C:Superfamily: (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology

Query Match 63.0%; Score 34; DB 2; Length 366;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9
 DB 289 EKVVPVGVD 297

Search completed: June 10, 2003, 13:49:14
 Job time : 11.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-12
Perfect score: 54
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	70.4	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	36	66.7	102	1 Y1LK_TVDVA	P31619 tobacco yel
3	36	66.7	1049	1 SPS_ORYSA	Q43802 oryza sativ
4	36	66.7	1068	1 SPS_MAIZE	P31927 zea mays (m
5	36	66.7	1081	1 SPS2_CRAPL	Q04933 craterostig
6	35	64.8	2747	1 FAF_DRONE	P55824 drosophila
7	34.5	63.9	748	1 KHLI_HUMAN	Q9nr64 homo sapien
8	34	63.0	154	1 PLAS_ORYSA	P20423 oryza sativ
9	34	63.0	155	1 PLAS_HORVU	P08248 hordeum vul
10	34	63.0	168	1 PLAT_POPNI	P11970 populus nig
11	34	63.0	566	1 SYEB_BORBU	P94283 borrelia bu
12	33	61.1	276	1 Y939_METJA	Q58349 methanococc
13	33	61.1	283	1 PANC_PSAE	Q9hv69 pseudomonas
14	33	61.1	394	1 HMPA_VIBCH	P08503 vibrio chol
15	33	61.1	421	1 ACBM_RAT	P08503 rattus norv
16	33	61.1	421	1 ECB2_HALEL	Q52250 halomonas e
17	33	61.1	423	1 ECB1_HALEL	Q52250 halomonas e
18	33	61.1	787	1 ECE2_HUMAN	O60344 homo sapien
19	33	61.1	801	1 FGR3_MOUSE	Q61851 mus musculu
20	33	61.1	806	1 CEK2_CHICK	P18460 gallus gall
21	33	61.1	877	1 SULH_SCHPO	O74377 schizosacch
22	33	61.1	982	1 ENV_SFV3L	P27399 simian foam
23	33	61.1	1401	1 RPOC_VIBCH	Q9kv29 vibrio chol
24	33	61.1	2717	1 ZEPI_HUMAN	P15822 homo sapien
25	32.5	60.2	472	1 ET2A_XENLA	P19102 xenopus lae
26	32	59.3	97	1 PLAS_DAUCA	P20422 daucus caro
27	32	59.3	175	1 HES3_RAT	Q04667 rattus norv
28	32	59.3	231	1 ARAD_ECOLI	P08203 escherichia
29	32	59.3	231	1 ARAD_SALTY	P06190 salmonella
30	32	59.3	233	1 HIS9_THEMA	Q9wzrl thermotoga
31	32	59.3	288	1 CGD2_RAT	Q04827 homo sapien
32	32	59.3	289	1 CGD2_HUMAN	P30279 homo sapien
33	32	59.3	289	1 CGD2_MOUSE	P30280 mus musculu

34	59.3	291	1 CGD1_BRARE	Q90459 brachydanio
35	59.3	291	1 CGD1_XENLA	P50755 xenopus lae
36	59.3	291	1 CGD2_CHICK	P49706 gallus gall
37	59.3	291	1 CGD2_XENLA	P53782 xenopus lae
38	59.3	292	1 CGD1_CHICK	P55169 gallus gall
39	59.3	292	1 CGD3_HUMAN	P30281 homo sapien
40	59.3	295	1 CGD1_HUMAN	P24385 homo sapien
41	59.3	295	1 CGD1_MOUSE	P25322 mus musculu
42	59.3	295	1 CGD1_RAT	P39948 rattus norv
43	59.3	338	1 MTBA_METBA	Q30640 methanosarc
44	59.3	472	1 ET2B_XENLA	Q91712 xenopus lae
45	59.3	561	1 HNFB_XENLA	Q91910 xenopus lae
46	59.3	759	1 SCTL_YEAST	P32784 saccharomyc
47	59.3	866	1 RECE_ECOLI	P00332 escherichia
48	59.3	995	1 HPI1_HUMAN	O00291 homo sapien
49	59.3	1176	1 NIR_NEUCR	P38661 neurospora
50	59.3	1258	1 ACN1_MOUSE	Q61137 mus musculu
51	59.3	1394	1 LTBS_HUMAN	P22084 homo sapien
52	59.3	1498	1 Y1A9_CLOAB	Q04351 clostridium
53	59.3	1595	1 LTBL_HUMAN	Q14766 homo sapien
54	59.3	1712	1 LTBL_RAT	Q00918 rattus norv
55	59.3	3174	1 CHAC_HUMAN	Q06r17 homo sapien
56	59.3	98	1 PLAS_ENTPR	P07485 enteromorph
57	59.3	98	1 PLAS_ULVPE	P13133 ulva arasak
58	59.3	98	1 PLAS_ULVPE	P56274 ulva pertus
59	59.3	99	1 PLAS_RUMOB	P00298 rumex obtus
60	59.3	99	1 PLAS_TOBAC	P23476 nicotiana t
61	59.3	124	1 REV_SIVCZ	P17280 chimpanzee
62	59.3	175	1 HES3_MOUSE	Q61657 mus musculu
63	59.3	258	1 YC10_METJA	Q58607 methanococc
64	59.3	319	1 YHAI_CRYPA	P10941 cryphonectr
65	59.3	327	1 YK14_CAEEL	P34338 caenorhabdi
66	59.3	338	1 MTBA_METAC	P58869 methanosarc
67	59.3	346	1 HYPE_BRAJA	P31906 bradyrhizob
68	59.3	391	1 LEI1_METTH	O27667 methanobact
69	59.3	427	1 TOLB_HAEIN	P44677 haemophilus
70	59.3	469	1 LET1_KLULA	P53998 kluyveromyc
71	59.3	529	1 ENP3_HUMAN	O75355 homo sapien
72	59.3	529	1 GUAA_MYLE	P46810 mycobacteri
73	59.3	625	1 GIDA_STAMM	Q99qt4 staphylococ
74	59.3	692	1 DNK1_SYNY3	Q55154 synechocyst
75	59.3	788	1 CY14_NEUCR	P23622 neurospora

ALIGNMENTS

RESULT 1				
CARB_FUSNN				
ID	CARB_FUSNN	STANDARD;	PRT;	1058 AA.
AC	Q8RG86;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).			
GN	CARB OR FN0422.			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteriia; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21986394; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyrpides N., Overbeek R.;			
RT	Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586.;			
RL	J. Bacteriol. 184:2005-2018(2002).			
CC	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			

```

RT      "The nucleotide sequence of the infectious cloned DNA component of
RT      tobacco yellow dwarf virus reveals features of geminiviruses
RT      infecting monocotyledonous plants.";
RT      Virology 187:633-642(1992).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M81103; AAA47947.1; -.
CC      PIR; A42452; A42452.
CC      InterPro; IPR002621; Gemini_mov.
CC      Pfam; PF01708; Gemini_mov; 1.
CC      Hypothetical protein.
CC      KW
CC      SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;
CC      -----
DR      QUERY Match      66.7%; Score 36; DB 1; Length 102;
DR      Best Local Similarity 60.0%; Pred. No. 1.7;
DR      Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DR      -----
QY      2 EVVPGMDYS 11
DB      :||| |:|:|
       7 QVPPSGINYS 16
DR      -----
RESULT 3
SPS_ORYSA
AC      SPS_ORYSA      STANDARD;      PRT; 1049 AA.
Q43802;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      sucrose-phosphate synthase (EC 2.4.1.14)
DE      (UDP-glucose-fructose-phosphate glucosyltransferase).
DE      Oryza sativa (Rice).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzaceae; Oryza.
OC      NCBI_TaxID=4530;
RR      [1]
RR      SEQUENCE FROM N.A.
RR      STRAIN=cv. Indica-IR36; TISSUE=Leaf;
RR      MEDLINE=96233138; PubMed=8666248;
RR      Valdez-Alarcon J.J., Ferrando M., Jimenez-Moralla B.,
RR      Herrera-Estrella L.;
RR      "Characterization of a rice sucrose-phosphate synthase-encoding
RR      gene.";
RR      Gene 170:217-222(1996).
CC      -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC      THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC      THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC      PHOTOASSIMILATES OUT OF THE LEAF.
CC      -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC      sucrose 6-phosphate.
CC      -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC      MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC      -!- PATHWAY: SUCROSE SYNTHESIS.
CC      -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
CC      -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC      ENZYME FUNCTION (BY SIMILARITY).
CC      -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/
CC      or send an email to license@isb-sib.ch).
CC      -----

```

```
CC ----- U33175; AAC49379.1; -
DR EMBL: U33175; AAC49379.1; -
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Transferase; Glycosyltransferase; Phosphorylation.
FT DOMAIN 22 29 POLY-GLY.
FT DOMAIN 695 698 POLY-GLU.
FT DOMAIN 775 779 POLY-ARG.
SQ SEQUENCE 1049 AA; 116455 MW; ED862E2819AA4B04 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1049;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 436 VIPPGMDFS 444

RESULT 4
SPS_MAIZE STANDARD; PRT; 1068 AA.
ID SPS_MAIZE
AC F31927;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase).
GN SPS.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 71-74; 206-212; 471-481 AND
RP 872-892.
RX MEDLINE-92338837; PubMed-1840396;
RA Strain=CV, PIONEER 3184; TISSUE=Leaf;
RA Worrell A.C., Bruneau J.-M., Summerfelt K., Boersig M., Voelker T.A.;
RT "Expression of a maize sucrose phosphate synthase in tomato alters
RT leaf carbohydrate partitioning."
RL Plant Cell 3:1121-1130(1991).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -!- PATHWAY: Sucrose synthesis.
CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
CC -!- DEVELOPMENTAL STAGE: GERMINATING SEEDS OR MATURE LEAVES.
CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL: Y11795; CAAY72491.1; -
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Transferase; Glycosyltransferase; Phosphorylation; Multigene family.
FT DOMAIN 245 248 POLY-SER.
FT DOMAIN 256 264 POLY-GLU.
FT DOMAIN 787 790 POLY-ARG.
SQ SEQUENCE 1081 AA; 120933 MW; DD142DC2F1A72900 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1081;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 445 VIPPGMDFS 453

RESULT 5
SPS2_CRAPL STANDARD; PRT; 1081 AA.
ID SPS2_CRAPL
AC O04933;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase 2).
GN SPS2.
OS Craterostigma plantagineum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Torenieae;
OC Craterostigma.
OX NCBI_TaxID=4153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97451773; PubMed-9306694;
RA Ingram J., Chandler J.W., Gallagher L., Salamini F., Bartels D.;
RT "Analysis of cDNA clones encoding sucrose-phosphate synthase in
RT relation to sugar interconversions associated with dehydration in the
RT resurrection plant Craterostigma plantagineum Hochst."
RL Plant Physiol. 115:113-121(1997).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -!- PATHWAY: Sucrose synthesis.
CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

RESULT 6
FAF_DROME STANDARD; PRT; 2747 AA.
AC P55824;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
GN FAF.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93202020; PubMed=1295747;
RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;
RT "The fat facets gene is required for Drosophila eye and embryo development."
RL development 116:985-1000(1992).
CC -1- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A ROLE IN COMPOUND EYE ASSEMBLY AND OOGENESIS RESPECTIVELY. IN THE LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR FUNCTION.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O = ubiquitin + a thiol.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC
CC EMBL; L04959; AAF01345.1; -
CC EMBL; L04958; AAF01346.1; -
CC MEROPS; C19.007; -
CC FlyBase; FBgn0005632; faf.
CC InterPro; IPR001394; UCH-2.
CC Pfam; PF00442; UCH-1; 1.
CC Pfam; PF00443; UCH-2; 1.
CC PROSITE; PS00972; UCH_2.1; 1.
CC PROSITE; PS00973; UCH_2.2; 1.
CC PROSITE; PS00974; UCH_2.3; 1.
CC PROSITE; PS00975; UCH_2.4; 1.
KW ubl conjugation pathway; Hydrolase; Thiol protease;
KW Developmental protein; Vision; Alternative splicing.
FT ACT_SITE 1677 1677 BY SIMILARITY.
FT ACT_SITE 1978 1978 BY SIMILARITY.
FT ACT_SITE 1986 1986 BY SIMILARITY.
FT VARSPLIC 2705 2747 KCRVILKKLVESKDEEDATSAVTAATTEVTTSPTAS
FT VARIANT 2725 2725 S -> T.
FT ORQQL -> VTRANNV (IN SHORT ISOFORM).
SQ SEQUENCE 2747 AA; 307954 MW; 1D97659F7A7B2ADE CRC64;
Query Match 64.8%; Score 35; DB 1; Length 2747;
Best Local Similarity 54.5%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ERVWPXGMDYS 11

Db 1394 EVIVPDGQDFS 1404
RESULT 7
KHL1_HUMAN STANDARD; PRT; 748 AA.
ID KHL1_HUMAN
AC Q9NR64; Q9P238; Q9H4X4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kelch-like protein 1.
DE KHL1 OR KIAA1490.
GN KHL1 OR KIAA1490.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347694; PubMed=10888605;
RA Koob M.D., Nemes J.P., Benzow K.A.;
RT "The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KHL1).";
RL Hum. Mol. Genet. 9:1543-1551(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
RN [3]
RP SEQUENCE OF 179-409 FROM N.A.
RA Kay M.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY PLAY A ROLE IN ORGANIZING THE ACTIN CYTOSKELETON OF THE BRAIN CELLS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF252283; AAF81719.1; -
CC EMBL; AF252279; AAF81716.1; -
CC EMBL; AB040923; BAA96014.1; ALT_INIT.
CC EMBL; AL353738; CAC16128.1; -
CC Genes; HGNC:6352; KHL1.
CC MIM; 605332; -
CC InterPro; IPR000210; BTB_POZ.
CC InterPro; IPR001798; Kelch.
CC Pfam; PF00651; BTB; 1.
CC Pfam; PF01344; Kelch; 6.
CC PRINTS; PR00501; KELCHREPEAT.
CC SMART; SM00225; BTB; 1.
CC PROSITE; PS50097; BTB; 1.
KW Cytoskeleton; Actin-binding; Repeat.
FT DOMAIN 43 88 SER-RICH.
FT DOMAIN 212 279 BTB.
FT REPEAT 460 506 KELCH 1.
FT REPEAT 507 553 KELCH 2.
FT REPEAT 555 600 KELCH 3.
FT REPEAT 601 647 KELCH 4.
FT REPEAT 649 700 KELCH 5.

FT REPEAT 701 747 KELCH 6.
SQ SEQUENCE 748 AA; 83680 MW; C11043D8282F9FF9 CRC64;

Query Match 63.9%; Score 34.5; DB 1; Length 748;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEVVPXGMDY 10
Db 127 EEVVP-GMDF 135

RESULT 8

PLAS_ORYSA
ID PLAS_ORYSA STANDARD; PRT; 154 AA.
AC P20423; Q9SBB8;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plastocyanin, chloroplast precursor.
GN PETE.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ilpoom; TISSUE=Leaf;
RA Lee J.-S.;
RT "Molecular cloning and characterization of plastocyanin precursor in rice."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-154.
RC STRAIN=cv. Japonica;
RX MEDLINE=89386623; PubMed=2780537;
RA Yano H., Kamo M., Tsugita A., Aso K., Nozu Y.;
RT "The amino acid sequence of plastocyanin from rice (Oryza sativa, subspecies japonica).";
RL Protein Seq. Data Anal. 2:385-389(1989).
CC -!- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.
CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.
CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AF093636; AAC78108.1; -;
DR PIR; S06105; S06105.
DR PIR; JT0352; JT0352.
DR HSSP; P00289; 2PCF.
DR InterPro; IPR000923; BlueCu_1.
DR Pfam; PF00127; copper-bind; 1.
DR PRINTS; PR00156; COPPERBLUE.
DR ProDom; PD001235; Copper_blue; 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
KW Transit peptide.
FT CHAIN 1 57 CHLOROPLAST.
FT CHAIN 58 154 PLASTOCYANIN.
FT DOMAIN 58 154 PLASTOCYANIN-LIKE.
FT METAL 94 94 COPPER (BY SIMILARITY).
FT METAL 139 139 COPPER (BY SIMILARITY).
FT METAL 142 142 COPPER (BY SIMILARITY).

FT METAL 147 147 COPPER (BY SIMILARITY).
SQ SEQUENCE 154 AA; 15577 MW; E45725D25B5F400D CRC64;

Query Match 63.0%; Score 34; DB 1; Length 154;
Best Local Similarity 54.5%; Pred. No. 6.8;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
Db 100 EDVPSGVDVS 110

RESULT 9

PLAS_HORVU
ID PLAS_HORVU STANDARD; PRT; 155 AA.
AC P08248;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plastocyanin, chloroplast precursor.
GN PETE.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bomi;
RA Nielsen O.S.; Gausling K.;
RT "The precursor of barley plastocyanin: sequence of cDNA clones and gene expression in different tissues.";
RL FEBS Lett. 225:159-162(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. NK 1558;
RX MEDLINE=94039081; PubMed=8223592;
RA Nielsen P., Gausling K.;
RT "In vitro binding of nuclear proteins to the barley plastocyanin gene promoter region.";
RL Eur. J. Biochem. 217:97-104(1993).
CC -!- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.
CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.
CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; Y00704; CAA86696.1; -;
DR EMBL; Z28347; CAA82201.1; -;
DR PIR; S00206; S00206.
DR HSSP; P00289; 2PCF.
DR InterPro; IPR000923; BlueCu_1.
DR Pfam; PF00127; copper-bind; 1.
DR PRINTS; PR00156; COPPERBLUE.
DR ProDom; PD001235; Copper_blue; 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
KW Transit peptide.
FT CHAIN 1 58 CHLOROPLAST.
FT CHAIN 59 155 PLASTOCYANIN.
FT DOMAIN 59 155 PLASTOCYANIN-LIKE.
FT METAL 95 95 COPPER (BY SIMILARITY).
FT METAL 140 140 COPPER (BY SIMILARITY).
FT METAL 143 143 COPPER (BY SIMILARITY).

FT METAL 148 148 COPPER (BY SIMILARITY).
 FT VARIANT 120 120 T -> N (IN CV. NK 1558).
 SQ SEQUENCE 155 AA; 15709 MW; DAA7EAB5F6F4F91 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 155;
 Best Local Similarity 54.5%; Pred. No. 6.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: ||| |
 Db 101 EDAPVSGVDVS 111

RESULT 10

PLAT_POPNI
 ID PLAT_POPNI STANDARD; PRT; 168 AA.
 AC P11970;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin B, chloroplast precursor.
 GN PETE.
 OS Populus nigra (lombardy poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Italica; TISSUP=Leaf;
 RA Reichert J., Jenzelewski V., Haehnel W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 70-168.
 RC STRAIN=cv. Italica;
 RA Dmitrov M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;
 RT "Complete amino acid sequence of poplar plastocyanin b.";
 RL FEBS Lett. 226:17-22(1987).
 CC -1- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF
 CC POPLAR PLASTOCYANINS A AND B.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; Z50186; CAA90565.1; -.
 FIR; S00210; S00210.
 HSSP; P00299; IPLC.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_Blue.
 DR Pfam; PF00127; copper_bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD001235; Copper_blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide; Multigene family.
 FT TRANSIT 1 69 CHLOROPLAST.
 FT CHAIN 70 168 PLASTOCYANIN B.
 FT DOMAIN 70 168 PLASTOCYANIN-LIKE.
 FT METAL 106 106 COPPER.
 FT METAL 153 153 COPPER.
 FT METAL 156 156 COPPER.
 FT METAL 161 161 COPPER.
 SQ SEQUENCE 168 AA; 16981 MW; F20DA6EA2038AEEA CRC64;

Query Match 63.0%; Score 34; DB 1; Length 168;
 Best Local Similarity 54.5%; Pred. No. 7.5;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 I: ||| |
 Db 112 EDAPVSGVDVS 122

RESULT 11

SYFB_BORBU
 ID SYFB_BORBU STANDARD; PRT; 566 AA.
 AC P94283;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
 DE tRNA ligase beta chain) (PHERS).
 GN PHET OR BB0514.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RA Barbour A.G., Hinnebusch J.;
 RT "phenylalanyl-tRNA synthetase genes (alpha and beta subunits) and
 RT thiorodoxin reductase gene of Borrelia burgdorferi.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 RT burgdorferi.";
 RL Nature 390:580-586(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
 CC diphosphate + L-phenylalanyl-tRNA(Phe).
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
 CC CHAIN FAMILY, SUBFAMILY 2.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; U82978; AAB41019.1; -.
 DR EMBL; AE001153; AAC66870.1; -.
 DR TIGR; BB0514; -.
 DR InterPro; IPR005147; B5.
 DR InterPro; IPR004531; PheT_arch.
 DR Pfam; PF03484; B5; 1.
 DR TIGRFAMS; TIGR00471; pheT_arch; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 566 AA; 65173 MW; 9D48C8B5D6D3B74B CRC64;

Query Match 63.0%; Score 34; DB 1; Length 566;


```

[1]
SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PA01;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
Nature 406:959-964(2000).
-!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
diphosphate + (R)-pantothenate.
-!- PATHWAY: Pantothenate biosynthesis; last step.
-!- SIMILARITY: BELONGS TO THE PANTOTHENATE SYNTHETASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@sib-sib.ch).
-----
EMBL: AE004886; AAC08116.1; -.
InterPro: IPR003721; Pantoate_ligase.
Pfam: PF02569; Pantoate_ligase; 1.
TIGRFAMs: TIGR00018; panc; 1.
Pantothenate biosynthesis; Ligase; Complete proteome.
SEQUENCE 283 AA; 30836 MW; C494949AB40E14E7 CRC64;
-----
Query Match 61.1%; Score 33; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9
Db ||: |||
96 EEMYPDGM 104

RESULT 14
HMPA_VIBCH STANDARD; PRT; 394 AA.
AC Q9KMY3;
AT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
DE HMP OR VCA0183.
GN Vibrio cholerae.
OS Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OC NCBI_TaxID=666;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN-EL Tor N16961 / Serotype O1;
RC MEDLINE-20406833; PubMed-10952301;
RX Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RC "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
CC -!- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS: ONE IS A HEME-CONTAINING
CC OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN
CC FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN
CC FLAVOHEMOPROTEINS SUBFAMILY.

```

CC CC -1- SIMILARITY: TO A NUMBER OF FAD/NAD(P) FLAVOPROTEIN
CC CC OXIDOREDUCTASES.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----

DR EMBL; AE004358; AAF96096.1; -
DR HSP; P39662; 1COX.
DR TIGR; VCA0183;
DR InterPro; IPR001834; Cyt_B5_reductase.
DR InterPro; IPR001709; FPN_cyt_reductse.
DR InterPro; IPR000971; Globin.
DR InterPro; IPR001433; Oxred_FAD/NAD(P).
DR InterPro; IPR001221; Phe_hydroxylase.
DR Pfam; PF00042; globin; 1.
DR Pfam; PF00175; NAD_binding; 1.
DR Pfam; PF00970; FAD_binding_6; 1.
DR PRINTS; PR00371; FPNCR.
DR PRINTS; PR00410; PHEHYDRXLASE.
DR PROSITE; PS01033; GLOBIN; 1.
KW Oxidoreductase; NADP; Heme; Flavoprotein; FAD; Iron transport;
KW Oxygen transport; Transport; Complete proteome.
FT DOMAIN 1 136
FT METAL 53 53
FT IRON (HEME DISTAL LIGAND)
FT IRON (HEME PROXIMAL LIGAND)
FT METAL 85 85
FT NP_BIND 268 273
FT SEQUENCE 394 AA; 44191 MW; DDA3490FAE28823A CRC64;
FT SQ

Query Match 61.1%; Score 33; DB 1; Length 394;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 EYVVPXGMDY 10
DB 194 EVTPEGSY 202
|||
|||

RESULT 15
ACDM_RAT ACDM_RAT STANDARD; PRT; 421 AA.
ID F08503;
AC 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
DE (EC 1.3.99.3) (MCAD).
GN ACADM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=87280028; PubMed=3611054;
RA Matsubara Y., Kraus J.P., Orasa H., Glassberg R., Finocchiaro G.,
RA Ikeda Y., Mole J., Rosenberg L.B., Tanaka K.;
RT "Molecular cloning and nucleotide sequence of cDNA encoding the
RT entire precursor of rat liver medium chain acyl coenzyme A
RT dehydrogenase."
RL J. Biol. Chem. 262:10104-10108(1987).
CC -1- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO
CC 16.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
CC ETF.
CC -1- COFACTOR: FAD.

CC CC -1- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
CC CC step.
CC CC -1- SUBUNIT: HOMOTETRAMER.
CC CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC CC -1- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
CC CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
CC CC TISSUES.
CC CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----

DR EMBL; J02791; AAA00670.1; -
DR PIR; A28436; DERTCM.
DR HSP; P11310; LEGD.
DR InterPro; IPR001552; Acyl-CoA_dh.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 25
FT CHAIN 26 421
FT ACYL-COA DEHYDROGENASE, MEDIUM-CHAIN
FT SPECIFIC.
FT ACT_SITE 193 193
FT FORMS A HYDROGEN-BOND WITH THE FLAVIN
FT N(5) OF THE FAD COFACTOR (BY SIMILARITY).
FT ACT_SITE 401 401
FT BASE (BY SIMILARITY).
FT SEQUENCE 421 AA; 46555 MW; 2CF076F8C919BDE8 CRC64;
FT SQ

Query Match 61.1%; Score 33; DB 1; Length 421;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
DB 58 EEIIPVADY 67
|||
|||

RESULT 16
ECB2_HALEL ECB2_HALEL STANDARD; PRT; 421 AA.
ID O52250;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46) (L-
DE diaminobutyric acid transaminase) (Diaminobutyrate transaminase).
GN ECTB.
OS Halomonas elongata.
OC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae;
OC Halomonas.
OX NCBI_TaxID=2746;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 2581T;
RX MEDLINE=98231640; PubMed=9570121;
RA Goller K., Ofer A., Galinski E.A.;
RT "Construction and characterization of an NaCl-sensitive mutant of
RT Halomonas elongata impaired in ectoine biosynthesis."
RL FEMS Microbiol. Lett. 161:293-300(1998).
CC -1- CATALYTIC ACTIVITY: L-2,4-diaminobutanoate + pyruvate = L-
CC aspartate 4-semialdehyde + L-alanine.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF ECTOINE (1,4,5,6-
CC TETRAHYDRO-2-METHYL-4-PYRIMIDINE CARBOXYLIC ACID).
CC -1- SUBUNIT: HOMOHXAMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF031489; AAC15882.1; -
 CC HSSP: P12995; 1QJ3.
 CC InterPro: IPR000954; Aminotran_3.
 CC InterPro: IPR004637; Dat.
 CC Pfam: PF00202; aminotran_3; 1.
 CC TIGRFAMs: TIGR00709; dat; 1.
 CC PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.
 CC Transfrase: Aminotransferase; Pyridoxal phosphate.
 CC BINDING 267 267 PYRIDOXAL PHOSPHATE (POTENTIAL).
 CC SEQUENCE 421 AA; 46166 MW; 44A2E21596E1E16C CRC64;
 CC -----
 CC Query Match 61.1%; Score 33; DB 1; Length 421;
 CC Best Local Similarity 58.3%; Pred. No. 32;
 CC Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
 CC
 CC QY 1 EEVW--PXGMDY 10
 CC |||: | | :||
 CC DB 91 EEVILKPRGLDY 102
 CC
 CC RESULT 17
 CC ECBI_HALEL STANDARD; PRT; 423 AA.
 CC ID ECBI_HALEL
 CC AC Q9ZEU7;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46) (L-
 CC de diaminobutyric acid transaminase) (Diaminobutyrate transaminase).
 CC GN ECTB.
 CC OS Halomonas elongata.
 CC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae;
 CC OC Halomonas.
 CC OX NCBI_TaxID=2746;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC STRAIN=DSM 3043;
 CC MEDLINE=99123891; PubMed=9924816;
 CC RA Canovas D., Vargas C., Calderon M.I., Ventosa A., Nieto J.J.;
 CC "Characterization of the genes for the biosynthesis of the compatible
 CC solute ectoine in the moderately halophilic bacterium Halomonas
 CC elongata DSM 3043.";
 CC RL Syst. Appl. Microbiol. 21:487-497(1998).
 CC -1- CATALYTIC ACTIVITY: L-2,4-diaminobutanate + pyruvate = L-
 CC aspartate 4-senialdehyde + L-alanine.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF ECTOINE (1,4,5,6-
 CC TETRAHYDRO-2-METHYL-4-PYRIMIDINE CARBOXYLIC ACID).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ011103; CAA09484.1; -
 CC HSSP: P12995; 1QJ3.
 CC InterPro: IPR000954; Aminotran_3.
 CC InterPro: IPR004637; Dat.
 CC Pfam: PF00202; aminotran_3; 1.
 CC TIGRFAMs: TIGR00709; dat; 1.
 CC PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.
 CC Transfrase: Aminotransferase; Pyridoxal phosphate.
 CC BINDING 267 267 PYRIDOXAL PHOSPHATE (POTENTIAL).
 CC SEQUENCE 421 AA; 46166 MW; 44A2E21596E1E16C CRC64;
 CC -----
 CC Query Match 61.1%; Score 33; DB 1; Length 421;
 CC Best Local Similarity 58.3%; Pred. No. 32;
 CC Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
 CC
 CC QY 1 EEVW--PXGMDY 10
 CC |||: | | :||
 CC DB 91 EEVILKPRGLDY 102
 CC
 CC RESULT 18
 CC ECE2_HUMAN STANDARD; PRT; 787 AA.
 CC ID ECE2_HUMAN
 CC AC O60344; Q96NX4; Q96NX3;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).
 CC GN ECE2 OR KIAA0604.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC [1]
 CC RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 CC RA Lorenzo M.-N., Khan R.Y., Wang Y., Tai S.C., Chan G.C., Cheung A.H.,
 CC Marsden P.A.;
 CC "Human endothelin converting enzyme-2 (ECE2): characterization of mRNA
 CC species and chromosomal localization.";
 CC RL Biochim. Biophys. Acta 0:0-0(2002).
 CC [2]
 CC RN SEQUENCE FROM N.A. (ISOFORM B).
 CC RC TISSUE=Brain;
 CC RX MEDLINE=98290545; PubMed=9628581;
 CC RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 CC Nomura N., Ohara O.;
 CC "Prediction of the coding sequences of unidentified human genes. IX.
 CC The complete sequences of 100 new cDNA clones from brain which can
 CC code for large proteins in vitro.";
 CC RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1 (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-
 CC Trp-1-Val-22 bond in the precursor.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; ECE-2A (shown here), ECE-2B and
 CC ECE-2C; are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF428263; AAL30386.1; -
 CC EMBL: AF428264; AAL30387.1; -
 CC EMBL: AF192531; AAG28399.1; -
 CC EMBL: AB011176; BAA25530.1; -
 CC HSSP: P08473; 1DWT.
 CC MEROPS: M13.003; -
 CC InterPro: IPR000718; Peptidase_M13.
 CC InterPro: IPR000130; Zn_MTpeptdse.

DR InterPro: IPR000954; Aminotran_3.
 DR InterPro: IPR004637; Dat.
 DR Pfam: PF00202; aminotran_3; 1.
 DR TIGRFAMs: TIGR00709; dat; 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.
 DR Transfrase: Aminotransferase; Pyridoxal phosphate.
 DR BINDING 267 267 PYRIDOXAL PHOSPHATE (POTENTIAL).
 DR SEQUENCE 423 AA; 46200 MW; 735C6BCF5A88288C CRC64;
 DR
 DR Query Match 61.1%; Score 33; DB 1; Length 423;
 DR Best Local Similarity 58.3%; Pred. No. 32;
 DR Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
 DR
 DR QY 1 EEVW--PXGMDY 10
 DR |||: | | :||
 DR DB 91 EEVILKPRGLDY 102
 DR
 DR RESULT 18
 DR ECE2_HUMAN STANDARD; PRT; 787 AA.
 DR ID ECE2_HUMAN
 DR AC O60344; Q96NX4; Q96NX3;
 DR DT 30-MAY-2000 (Rel. 39, Created)
 DR DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DR DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DR DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).
 DR GN ECE2 OR KIAA0604.
 DR OS Homo sapiens (Human).
 DR OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 DR OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 DR OX NCBI_TaxID=9606;
 DR [1]
 DR RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 DR RA Lorenzo M.-N., Khan R.Y., Wang Y., Tai S.C., Chan G.C., Cheung A.H.,
 DR Marsden P.A.;
 DR "Human endothelin converting enzyme-2 (ECE2): characterization of mRNA
 DR species and chromosomal localization.";
 DR RL Biochim. Biophys. Acta 0:0-0(2002).
 DR [2]
 DR RN SEQUENCE FROM N.A. (ISOFORM B).
 DR RC TISSUE=Brain;
 DR RX MEDLINE=98290545; PubMed=9628581;
 DR RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 DR Nomura N., Ohara O.;
 DR "Prediction of the coding sequences of unidentified human genes. IX.
 DR The complete sequences of 100 new cDNA clones from brain which can
 DR code for large proteins in vitro.";
 DR RL DNA Res. 5:31-39(1998).
 DR -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1 (BY
 DR SIMILARITY).
 DR -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-
 DR Trp-1-Val-22 bond in the precursor.
 DR -1- SUBCELLULAR LOCATION: Type II membrane protein.
 DR -1- ALTERNATIVE PRODUCTS: 3 isoforms; ECE-2A (shown here), ECE-2B and
 DR ECE-2C; are produced by alternative splicing.
 DR -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13.
 DR -----
 DR This SWISS-PROT entry is copyright. It is produced through a collaboration
 DR between the Swiss Institute of Bioinformatics and the EMBL outstation -
 DR the European Bioinformatics Institute. There are no restrictions on its
 DR use by non-profit institutions as long as its content is in no way
 DR modified and this statement is not removed. Usage by and for commercial
 DR entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 DR or send an email to license@isb-sib.ch).
 DR -----
 DR EMBL: AF428263; AAL30386.1; -
 DR EMBL: AF428264; AAL30387.1; -
 DR EMBL: AF192531; AAG28399.1; -
 DR EMBL: AB011176; BAA25530.1; -
 DR HSSP: P08473; 1DWT.
 DR MEROPS: M13.003; -
 DR InterPro: IPR000718; Peptidase_M13.
 DR InterPro: IPR000130; Zn_MTpeptdse.

FT CARBOHYD 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 305 TAGANTDDKEVLSLHVTFEDAGEYTCLAGNSIGFESHHS
 FT AWLVVLP -> SWISENVEADARLRLANVSERDGGEYLCRA
 FT TNFIGNVAEKAFWLKRVHPQA (IN ISOFORM 2).
 FT P -> L (IN REF. 2).
 FT CONFLICT 684 MISSING (IN REF. 2).
 FT CONFLICT 687 MISSING (IN REF. 2).
 SQ SEQUENCE 801 AA; 87758 MW; 68BC110212691705 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 801;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PGMGMDYS 11
 | | | | |
 DB 566 PGMGMDYS 572

RESULT 20
 CEK2_CHICK STANDARD; PRT; 806 AA.
 AC P18450;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine kinase receptor CEK2 precursor (EC 2.7.1.112).
 GN CEK2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90332672; PubMed=2165604;
 RA Pasquale E.B.;
 RT "A distinctive family of embryonic protein-tyrosine kinase
 receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; M35195; AAA48664.1; .
 DR PIR; A35963; A35963.
 DR HSP; P11362; 1FKG.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig; 3.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00408; IGG2; 3.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
 KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
 KW Repeat; Signal.

FT SIGNAL 1 19
 FT CHAIN 20 806
 FT DOMAIN 20 364
 FT TRANSMEM 365 389
 FT DOMAIN 390 806
 FT DOMAIN 54 114
 FT DOMAIN 163 229
 FT DOMAIN 262 340
 FT DOMAIN 131 141
 FT DOMAIN 466 755
 FT NP_BIND 472 480
 FT BINDING 502 502
 FT ACT_SITE 611 611
 FT MOD_RES 642 642
 FT DISULFID 61 107
 FT DISULFID 170 222
 FT DISULFID 269 333
 FT CARBOHYD 83 83
 FT CARBOHYD 96 96
 FT CARBOHYD 118 118
 FT CARBOHYD 219 219
 FT CARBOHYD 256 256
 FT CARBOHYD 288 288
 FT CARBOHYD 309 309
 FT CARBOHYD 322 322
 SQ SEQUENCE 806 AA; 89730 MW; B38B3C6D5F2314B6 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 806;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PGMGMDYS 11
 | | | | |
 DB 566 PGMGMDYS 572

RESULT 21
 SULH_SCHPO STANDARD; PRT; 877 AA.
 AC Q74377;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable sulfate permease C3H7.02.
 GN SPBC3H7.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odele C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti I., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "the genome sequence of Schizosaccharomyces pombe.";
CC Nature 415:871-880(2002).
CC -1- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL031261; CAA20298.1; -
CC InterPro; IPR002645; STAS.
CC Pfam; PF00916; Sulfate_transp; 1.
CC Pfam; PF01740; STAS; 1.
CC TIGRFAMS; TIGR00815; sulp; 1.
CC PROSITE; PS01130; SLC26A; 1.
CC PROSITE; PS0801; STAS; 1.
CC Transport; Transmembrane.
CC TRANSMEM 133 153 POTENTIAL.
CC TRANSMEM 161 181 POTENTIAL.
CC TRANSMEM 186 206 POTENTIAL.
CC TRANSMEM 221 241 POTENTIAL.
CC TRANSMEM 243 263 POTENTIAL.
CC TRANSMEM 292 312 POTENTIAL.
CC TRANSMEM 329 349 POTENTIAL.
CC TRANSMEM 384 404 POTENTIAL.
CC TRANSMEM 424 444 POTENTIAL.
CC TRANSMEM 461 481 POTENTIAL.
CC TRANSMEM 484 504 POTENTIAL.
CC TRANSMEM 518 538 POTENTIAL.
CC TRANSMEM 543 563 POTENTIAL.
CC DOMAIN 594 747 STAS.
CC SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;

CC Query Match 61.1%; Score 33; DB 1; Length 877;
CC Best Local Similarity 66.7%; Pred. No. 70;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CC QY 3 VVPXGMDYS 11
CC ||| || |:
CC Db 148 VVPQGSYA 156

CC RESULT 22
CC ENV_SFV3L
CC ID ENV_SFV3L STANDARD; PRT; 982 AA.
CC AC P27399;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
CC DE ENV polyprotein (Coat polyprotein).
CC GN ENV.
CC OS Simian foamy virus (type 3 / strain LK3) (SFV-3).
CC OC Viruses; Retroviridae; Retroviridae; Spumavirus.
CC OX NCBI_TaxID=11644;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=92124734; PubMed=1310187;
CC RA Renne R., Friedl E., Schweizer M., Fleps U., Turek R.,
CC RA Neumann-Haefelin D.;
CC RT "Genomic organization and expression of simian foamy virus type 3

(SFV-3).";
RT Virology 186:597-608(1992).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74895; AAA47798.1; ALT_INIT.
CC PIR; C40820; VCLJLK.
CC InterPro; IPR005070; Foamy_virus_1.
CC Pfam; PF03408; Foamy_virus_1.
CC Coat protein; Transmembrane; Glycoprotein; Glycoprotein.
CC TRANSMEM 68 88 I (POTENTIAL).
CC TRANSMEM 95 975 II (POTENTIAL).
CC CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 346 346 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 484 484 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 551 551 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 776 776 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 982 AA; 113313 MW; 721F2F8929D604FF CRC64;

CC Query Match 61.1%; Score 33; DB 1; Length 982;
CC Best Local Similarity 66.7%; Pred. No. 79;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CC QY 1 EEVVPXGMD 9
CC |||| | |
CC Db 44 EEVPTPRMD 52

CC RESULT 23
CC RPOC_VIBCH
CC ID RPOC_VIBCH STANDARD; PRT; 1401 AA.
CC AC Q9KV29;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
CC DE beta' chain) (RNA polymerase beta' subunit).
CC GN RPOC OR VC0329.
CC OS Vibrio cholerae.
CC OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
CC OX NCBI_TaxID=666;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=El Tor N16961 / Serotype O1;
CC RX MEDLINE=20406833; PubMed=10952301;
CC RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
CC RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
CC RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
CC RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
CC RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
CC RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
CC RA Fraser C.M.;
CC RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
CC RT cholerae.";
CC RL Nature 406:477-483(2000).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES (By similarity).
CC -! CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -! SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN (By similarity).
CC -! SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE004121; AAF93502.1; -.
CC DR HSP: O9KWD6; 1HQW.
CC DR TIGR: VC0329; -.
CC DR InterPro: IPR000722; RNA_pol_A.
CC DR InterPro: IPR002879; RNA_pol_A2.
CC DR Pfam: PF00623; RNA_pol_A; 1.
CC DR Pfam: PF01854; RNA_pol_A2; 2.
CC DR Transferase: DNA-directed RNA polymerase; Transcription;
CC KW Complete proteome
CC SQ SEQUENCE 1401 AA; 155021 MW; DFDB0F2B5514504F CRC64;
CC -----
CC Query Match 61.1%; Score 33; DB 1; Length 1401;
CC Best Local Similarity 50.0%; Pred. No. 1.2e+02;
CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 2 EVVPXGMDYS 11
CC DB 581 QIVPKGLPYS 590
CC -----
CC RESULT 24
CC ZEP1_HUMAN STANDARD; PRT; 2717 AA.
CC ID ZEP1_HUMAN
CC AC P15822;
CC DT 01-APR-1990 (Rel. 14, Created)
CC DT 01-APR-1990 (Rel. 14, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Zinc finger protein 40 (human immunodeficiency virus type I enhancer-
CC binding protein 1) (HIV-EPI) (Major histocompatibility complex binding
CC protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
CC (PRDII-BF1).
CC GN HIVEP1 OR ZNF40.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=90169514; PubMed=2106471;
CC RA Fan C.M., Maniatis T.;
CC RT "A DNA-binding protein containing two widely separated zinc finger
CC motifs that recognize the same DNA sequence.";
CC RL Genes Dev. 4:29-42(1990).
CC RN [2]
CC RP STRUCTURE BY NMR OF 2113-2142.
CC RX MEDLINE=91064333; PubMed=2248949;
CC RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
CC RA Gronenborn A.M.;
CC RT "High-resolution three-dimensional structure of a single zinc finger
CC from a human enhancer binding protein in solution.";
CC RL Biochemistry 29:9324-9334(1990).
CC RN [3]
CC RP STRUCTURE BY NMR OF 2087-2142.
CC RX MEDLINE=92232684; PubMed=1567844;
CC RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,
CC RA Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc
RT finger from the human enhancer binding protein MBP-1.";
RL Biochemistry 31:3907-3917(1992).
CC -! FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
CC '5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
CC IN T-CELL ACTIVATION.
CC -! SUBCELLULAR LOCATION: Nuclear.
CC -! INDUCTION: BY MITOGEN AND PHORBOL ESTER.
CC -! DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
CC ZINC-FINGER IN-BETWEEN.
CC -! SIMILARITY: STRONG, TO HIVEP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51435; CAA35798.1; -.
CC DR PIR: A34203; A34203
CC DR PDB: 3ZNF; 15-JAN-92.
CC DR PDB: 4ZNF; 15-JAN-92.
CC DR PDB: 1BBO; 31-OCT-93.
CC DR TRANSFAC: T00497; -.
CC DR Genew: HGNC:4920; HIVEP1.
CC DR MIM: 194540; -.
CC DR InterPro: IPR000822; Znf_C2H2.
CC DR Pfam: PF00096; Zf-C2H2; 5.
CC DR PRINTS: PR00048; ZINC_FINGER.
CC DR SMART: SM00355; Znf_C2H2; 4.
CC DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
CC DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 4.
CC KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
CC KW Nuclear protein; Repeat; 3D-structure.
CC FT DOMAIN 406 456
CC FT ZN_FING 406 428
CC FT ZN_FING 434 456
CC FT ZN_FING 803 806
CC FT DOMAIN 958 981
CC FT DOMAIN 2087 2139
CC FT ZN_FING 2115 2139
CC FT ZN_FING 2115 2139
CC FT STRAND 2088 2088
CC FT TURN 2090 2092
CC FT STRAND 2095 2095
CC FT HELIX 2099 2108
CC FT TURN 2109 2109
CC FT STRAND 2115 2116
CC FT STRAND 2123 2124
CC FT HELIX 2127 2135
CC SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;
CC -----
CC Query Match 61.1%; Score 33; DB 1; Length 2717;
CC Best Local Similarity 66.7%; Pred. No. 2.3e+02;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 3 VVPXGMDYS 11
CC DB 2405 VVPAGLTYS 2413
CC -----
CC RESULT 25
CC ET2A_XENLA STANDARD; PRT; 472 AA.
CC ID ET2A_XENLA
CC AC P19102;
CC DT 01-NOV-1990 (Rel. 16, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-ETS-2A protein.
GN ETS2A OR ETS-2A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=92158632; PubMed=1741266;
RA Burdett L.A., Qi S.W., Chen Z.Q., Lautenberger J.A., Papas T.S.;
RT "Characterization of the cDNA sequences of two xenopus ets-2 proto-
oncogenes";
RL Nucleic Acids Res. 20:371-371(1992).
RN [2]
RN SEQUENCE OF 121-472 FROM N.A.
RP TISSUE=Oocyte;
RX MEDLINE=90356411; PubMed=2201951;
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,
RA Stehelin D., Befort N., Remy P.;
RT "Isolation of two different c-ets-2 proto-oncogenes in Xenopus
laevis";
RL Nucleic Acids Res. 18:4603-4604(1990).
RN [3]
RN SEQUENCE OF 121-472 FROM N.A.
RX MEDLINE=92088972; PubMed=1751411;
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,
RA Stehelin D., Befort N., Remy P.;
RT "Cloning, sequencing, and expression of two Xenopus laevis c-ets-2
proto-oncogenes";
RL Cell Growth Differ. 2:447-456(1991).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81683; AAA49705.1; -;
DR EMBL; X51826; CAA36124.1; -;
DR PIR; S10994; S10994.
DR PIR; S28824; S28824.
DR HSP; P14921; 2STT.
DR TRANSFAC; T02041; -;
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR InterPro; IPR003118; SAM_PNT.
DR Pfam; PF00178; Ets; 1.
DR Pfam; PF02198; SAM_PNT; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR SMART; SM00251; SAM_PNT; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW DNA-binding; Nuclear protein.
FT DOMAIN 87 170 POINTED.
FT DNA_BIND 366 446 ETS-DOMAIN.
SQ SEQUENCE 472 AA; 53894 MW; E0E808B5E6BF111 CRC64;

Query Match 60.2%; Score 32.5; DB 1; Length 472;
Best Local Similarity 58.3%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 EEVPEXGMD-YS 11
|: |||: ||

Db 49 EQAVPTGLDSYS 60

Search completed: June 10, 2003, 13:40:21
Job time : 4.5 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time-25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-12
Perfect score: 54
Sequence: 1 EHVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SPTREMBL21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	74.1	156	3 Q12479	Q12479 saccharomyc
2	38	70.4	363	17 Q30260	Q30260 archaeoglob
3	38	70.4	1063	16 Q8RG86	Q8RG86 fusobacteri
4	36	66.7	341	10 Q22081	Q22081 citrus unsh
5	36	66.7	348	10 Q22096	Q22096 citrus unsh
6	36	66.7	452	10 Q8W568	Q8W568 arabidopsis
7	36	66.7	460	10 Q9C9T7	Q9C9T7 arabidopsis
8	36	66.7	1047	10 P93782	P93782 saccharum o
9	36	66.7	1083	10 Q9SN30	Q9SN30 arabidopsis
10	36	66.7	1084	10 Q43010	Q43010 oryza sativ
11	36	66.7	1100	10 Q8S064	Q8S064 oryza sativ
12	35	64.8	219	5 Q8GQ04	Q8GQ04 eriocheir s
13	35	64.8	253	16 Q8XPA8	Q8XPA8 clostridium
14	35	64.8	298	2 Q52367	Q52367 rhizobium t
15	35	64.8	425	5 Q9XVK4	Q9XVK4 caenorhabdi
16	35	64.8	433	16 Q9A382	Q9A382 caulobacter

RESULT 1

Q12479 ID Q12479
AC Q12479
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.

ALIGNMENTS

PRELIMINARY; PRT; 156 AA.

17	35	64.8	440	17	Q9YFI3	Q9yfi3 aeropyrum p
18	35	64.8	511	2	O52680	O52680 escherichia
19	35	64.8	517	16	O8XZL5	O8xzl5 ralstonia s
20	35	64.8	745	5	O95P46	O95p46 carcinus ma
21	35	64.8	1031	5	O9U6A3	O9u6a3 callinectes
22	35	64.8	1150	5	O17704	O17704 caenorhabdl
23	35	64.8	1410	2	O52673	O52673 escherichia
24	35	64.8	1420	2	O52666	O52666 escherichia
25	35	64.8	1474	17	O27146	O27146 methanobact
26	35	64.8	1828	16	O98K29	O98k29 rhizobium l
27	35	64.8	2778	5	O9V9T6	O9v9t6 drosophila
28	34	63.9	748	4	O8TBJ7	O8tbj7 homo sapien
29	34	63.0	143	5	O9VSY8	O9vsy8 drosophila
30	34	63.0	154	10	O9SBB8	O9sbb8 oryza sativ
31	34	63.0	215	16	O8R9L5	O8r9l5 thermoanaer
32	34	63.0	290	16	O8U7J0	O8u7j0 agrobacteri
33	34	63.0	296	17	O9YET8	O9yet8 aeropyrum p
34	34	63.0	357	17	O29920	O29920 archaeoglob
35	34	63.0	366	17	O29451	O29451 archaeoglob
36	34	63.0	387	16	O98FX1	O98fx1 rhizobium l
37	34	63.0	543	3	O8TFF4	O8tff4 trichoderma
38	34	63.0	558	16	O8R8Z2	O8r8z2 thermoanaer
39	34	63.0	565	16	O9CIN1	O9cin1 lactococcus
40	34	63.0	587	16	O9JZP8	O9jzp8 neisseria m
41	34	63.0	906	10	O9ZSY4	O9zsy4 arabidopsis
42	34	63.0	908	10	O9FJK8	O9fjk8 arabidopsis
43	34	63.0	908	10	O8W4J9	O8w4j9 arabidopsis
44	34	63.0	908	10	O9ZSY3	O9zsy3 arabidopsis
45	34	63.0	909	10	O9MSA1	O9msa1 arabidopsis
46	34	63.0	1062	2	P95422	P95422 pseudomonas
47	34	63.0	1062	16	O910Y8	O910y8 pseudomonas
48	34	63.0	3472	1	O74056	O74056 cenarchaeum
49	33	61.1	78	6	O9XST4	O9xst4 canis famil
50	33	61.1	97	17	O97VR9	O97vr9 sulfolobus
51	33	61.1	128	17	O97US8	O97us8 sulfolobus
52	33	61.1	172	13	O02528	O02528 oryzias lat
53	33	61.1	175	10	O8VY88	O8vy88 arabidopsis
54	33	61.1	183	2	O9S1I0	O9s1i0 sweet potat
55	33	61.1	184	17	O97XJ2	O97xj2 sulfolobus
56	33	61.1	193	5	O8SWL4	O8swl4 encephalito
57	33	61.1	209	13	O9DE16	O9de16 brachydanio
58	33	61.1	210	10	O6S890	O6s890 cyclotella
59	33	61.1	217	4	O00404	O00404 homo sapien
60	33	61.1	225	10	O40129	O40129 lycopersico
61	33	61.1	230	17	O9P9L6	O9p9l6 pyrobaculum
62	33	61.1	247	16	O92U66	O92u66 rhizobium m
63	33	61.1	257	10	O9C6J0	O9c6j0 arabidopsis
64	33	61.1	262	17	O97YB8	O97yb8 sulfolobus
65	33	61.1	267	17	O97Y57	O97y57 sulfolobus
66	33	61.1	299	4	O9UEE9	O9uee9 homo sapien
67	33	61.1	299	17	O97TV4	O97tv4 sulfolobus
68	33	61.1	299	17	O9UWN4	O9uwn4 sulfolobus
69	33	61.1	307	17	O9HP60	O9hpf60 halobacteri
70	33	61.1	396	17	O8TIT9	O8tit9 methanosarc
71	33	61.1	441	17	O9HIW6	O9hiw6 thermoplasma
72	33	61.1	556	4	O43733	O43733 homo sapien
73	33	61.1	577	15	O9QBR8	O9qbr8 simian foam
74	33	61.1	577	15	O9QBR7	O9qbr7 simian foam
75	33	61.1	577	15	O9QBR6	O9qbr6 simian foam

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA De haan M., Grivell L.A., Maarse A.C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-FY1679;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA STRAIN-FY1679;
 RL MEDLINE-94019318; PubMed-8413243;
 RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
 RT Sherman F.;
 RT "CYC2 encodes a factor involved in mitochondrial import of yeast
 cytochrome c.";
 RL Mol. Cell. Biol. 13:6442-6451(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA STRAIN-FY1679;
 RL MEDLINE-94169519; PubMed-7764548;
 RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
 RT "Molecular cloning of a gene, DHS1, which complements a drug-
 RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
 RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
 DR EMBL; Z74920; CAA99201.1;
 DR EMBL; X87331; CAA60762.1;
 DR SGD; S0005539; YOR013W.
 SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;
 Query Match 74.1%; Score 40; DB 3; Length 156;
 Best Local Similarity 77.8%; Pred. No. 1.7;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EVVFXGMDY 10
 Db ||| ||||
 50 EVNPLGNDY 58
 RESULT 2
 O30260
 ID O30260 PRELIMINARY; PRT; 363 AA.
 AC O30260;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein AF2411.
 GN AF2411.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE-98049343; PubMed-9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerkvliet A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91255.1;
 DR TIGR; AF2411;
 DR InterPro; IPR002103; Bac_luciferase.
 DR Pfam; PF00296; bac_luciferase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 363 AA; 41736 MW; 0E976EAE788F4803 CRC64;
 Query Match 70.4%; Score 38; DB 17; Length 363;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EVVFXGMDYS 11
 Db |::| |::|
 120 ENIVPYGIDFS 130
 RESULT 3
 O8RG86
 ID O8RG86 PRELIMINARY; PRT; 1063 AA.
 AC O8RG86;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
 GN FN0422.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 25586;
 RX MEDLINE-21886394; PubMed-11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010554; AAL94625.1;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;
 Query Match 70.4%; Score 38; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EVVFXGMDYS 11
 Db |::| |::|
 195 EIVPGLNYS 204
 RESULT 4
 O22081
 ID O22081 PRELIMINARY; PRT; 341 AA.
 AC O22081;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Sucrose-phosphate synthase (Fragment).
 GN CIMSFS2.
 OS Citrus unshiu (Satsuma orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=55188;

```

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV, MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RC MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
RT phosphate synthase isoforms from a citrus fruit (Citrus unshiu
RT Marc.).";
RL Mol. Gen. Genet. 252:346-351(1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=CV, MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RC Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR EMBL: AB006319; BAA23215.1; -.
FT NON_TER 341 341
FT SEQUENCE 341 AA; 38136 MW; 61417A69C4560777 CRC64;
Query Match 66.7%; Score 36; DB 10; Length 341;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMDYS 11
Db 228 VIPPGMDFS 236
RESULT 5
OZ2096 PRELIMINARY; PRT; 348 AA.
ID OZ2096;
AC OZ2096;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Sucrose-phosphate synthase (Fragment).
GN CITSPS3.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RC Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR EMBL: AB006660; BAA22071.1; -.
FT NON_TER 348 348
FT SEQUENCE 348 AA; 38556 MW; E81C21EBA6FF5C5E CRC64;
Query Match 66.7%; Score 36; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMDYS 11
Db 234 VIPPGMDFS 242

```

```

RESULT 6
Q8W568 PRELIMINARY; PRT; 452 AA.
ID Q8W568;
AC Q8W568;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Atlg73750/F25P22_17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.;
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF419606; AAL31938.1; -.
DR InterPro: IPR002471; Prol-endopep_ser.
DR InterPro: IPR000379; Ser-estrus_site.
DR PROSITE: PS00708; PRO-ENDOPEP_SER; UNKNOWN_1.
SQ SEQUENCE 452 AA; 49682 MW; A159955B21742C4A CRC64;
Query Match 66.7%; Score 36; DB 10; Length 452;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
Db 210 EEDVPSAMDY 219
[1]
RESULT 7
Q9C9T7 PRELIMINARY; PRT; 460 AA.
ID Q9C9T7;
AC Q9C9T7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 50.6 kDa protein.
GN F25P22.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV, COLUMBIA;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huitzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.-L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

```

```

RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL: AC012679; AAG52073.1; -.
DR InterPro: IPR002471; Prol_endopep_ser.
DR InterPro: IPR000379; Ser_estrs_site.
DR PROSITE: PS00708; PRO_ENDOPEP_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 50564 MW; E94B27B5C4B249EC CRC64;

Query Match 66.7%; Score 36; DB 10; Length 460;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
Db 218 EEDVPSAMDY 227
|| || |||
|| || |||

RESULT 8
P93782 PRELIMINARY; PRT; 1047 AA.
AC P93782;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (Fragment).
GN SOSPS1.
OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Sugiharto B., Sakakibara H., Sugiyama T.;
RT "Differential Expression of Two Genes for Sucrose-Phosphate Synthase
RT in Sugarcane: Molecular Cloning of the cDNAs and Comparative Analysis
RT of Gene Expression.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB001337; BAA19241.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 1047 AA; 116379 MW; D0EDB34961E1D83D CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1047;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 414 VIPPGMDFS 422
|| || |||
|| || |||

RESULT 9
Q9SN30 PRELIMINARY; PRT; 1083 AA.
AC Q9SN30;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Sucrose-phosphate synthase-like protein (EC 2.4.1.14).
GN F28M11.40 OR AT4G10120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```

```

RN SEQUENCE FROM N.A.
RP Bavan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049487; CAB39764.1; -.
DR EMBL: AL161516; CAB78135.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1083 AA; 122688 MW; EFC126FCA2137BB1 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 483 VIPPGMDFS 491
|| || |||
|| || |||

RESULT 10
Q43010 PRELIMINARY; PRT; 1084 AA.
AC Q43010;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).
GN SP51.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. JAPONICA;
RA Sakamoto M., Satozawa T., Kishimoto N., Higo K., Shimada H.,
RA Fujimura T.;
RT "Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS)
RT gene that is specifically expressed in the source organ.";
RL Plant Sci. 112:207-217(1995).
CC -1- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.
CC -1- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
DR EMBL: D45890; BAA08304.1; -.
DR InterPro: IPR001296; Glycos_transf_1.
DR InterPro: IPR001557; L_LDH.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR PROSITE: PS00064; L_LDH; 1.
KW Glycolysis; NAD; Oxidoreductase.
SQ SEQUENCE 1084 AA; 119417 MW; 4D0D79AFEDD2F92B CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
Db 453 VIPPGMDFS 461
|| || |||
|| || |||

RESULT 11
Q8S064

```

ID Q8S064 PRELIMINARY; PRT; 1100 AA.
 AC Q8S064;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative sucrose-phosphate synthase.
 GN P0678F11.14.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0678F11.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003437; BAB86107.1; -;
 SQ SEQUENCE 1100 AA; 121170 MW; 084F4604BA389CAD CRC64;
 Query Match 66.7%; Score 36; DB 10; Length 1100;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 VVPXGMDYS 11
 DB 469 VIPPGMDFS 477

RESULT 12
 Q9GQ04 PRELIMINARY; PRT; 219 AA.
 ID Q9GQ04;
 AC Q9GQ04;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Na+/K+/2Cl-cotransporter (Fragment).
 OS Eriocheir sinensis (Chinese mitten crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Grapsoidae; Varunidae; Eriocheir.
 OX NCBI_TaxID=95602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GILL;
 RA Wehrhau D., Towle D.W.;
 RT "Na+/H+-exchanger and Na+/K+/2Cl- cotransporter are expressed in
 RT gills of the euryhaline Chinese crab Eriocheir sinensis.";
 RL Comp. Biochem. Physiol. 126:S158-S158(2000).
 DR EMBL: AF301160; AKG39938.1; -;
 FT NON_TER 1
 FT NON_TER 219
 SQ SEQUENCE 219 AA; 24159 MW; 599442DA26F0D3F1 CRC64;
 Query Match 64.8%; Score 35; DB 5; Length 219;
 Best Local Similarity 75.08%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMDYS 11
 DB 107 VPOGLDYS 114

RESULT 13
 Q8XPA8 PRELIMINARY; PRT; 253 AA.
 ID Q8XPA8
 AC Q8XPA8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein CPE0057.

GN CPE0057.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 EX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yanashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kunara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL: AP003185; BAB79763.1; -;
 DR InterPro; IPR000205; NAD_binding.
 DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF00899; Thif; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 253 AA; 27714 MW; 814DF79D3E0D7486 CRC64;
 Query Match 64.8%; Score 35; DB 16; Length 253;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDY 10
 DB 108 EEIIPDDVDY 117

RESULT 14
 Q52367 PRELIMINARY; PRT; 298 AA.
 ID Q52367;
 AC Q52367;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Aryl-alcohol dehydrogenase homolog (Fragment).
 GN XYLBI.
 OS Rhizobium tropici.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=398;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CFN299;
 RA Rosenblueth M., Hynes M.F., Martinez-Romero E.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- COFACTOR: ZINC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 DR EMBL: AF036920; AAC04779.1; -;
 DR HSSP; P07846; ISDG.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000205; NAD_binding.
 DR Pfam; PF00107; adh_zinc; 1.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Plasmid; Zinc.
 FT NON_TER 298
 FT NON_TER 298
 SQ SEQUENCE 298 AA; 31092 MW; 49B2F8117C33AE87 CRC64;
 Query Match 64.8%; Score 35; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMDYS 11
 DB 250 EEIPEGADFS 259

RESULT 15

Q9XVK4
ID Q9XVK4 PRELIMINARY; PRT; 425 AA.
AC Q9XVK4
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE R10D12.10 protein.
GN R10D12.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Percy C.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z81109; CAB03241.1;
DR InterPro: IPR000719; Euk_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferrase.
SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 425;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
Db 335 EQIVPGLOY 344
:::|:|

RESULT 16
Q9A382
ID Q9A382 PRELIMINARY; PRT; 433 AA.
AC Q9A382
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Peptidoglycan-binding protein, putative.
GN CC3322.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-ATCC 19089 / CB15;
RA Nieman W.C.; Feldblyum T.V.; Laub M.T.; Paulsen I.T.; Nelson K.E.;
RA Eisen J.; Heidelberg J.F.; Alley M.R.K.; Ohta N.; Maddock J.R.;
RA Potocka I.; Nelson W.C.; Newton A.; Stephens C.; Phadke N.D.; Ely B.;
RA DeBoy R.T.; Dodson R.J.; Durkin A.S.; Gwinn M.L.; Haft D.H.;
RA Kolonay J.F.; Smit J.; Craven M.B.; Khouri H.; Shetty J.; Berry K.;
RA Utterback T.; Tran K.; Wolf A.; Vamathevan J.; Ermolaeva M.; White O.;
RA Salzberg S.L.; Venter J.C.; Shapiro L.; Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005994; AAK25284.1;
DR HSSP: P41052; 1LTM.
DR TIGR: CC3322;
DR InterPro: IPR002477; PG_binding.
DR Pfam: PF01471; PG_binding_1; 1.
KW Complete proteome.
SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 433;
Best Local Similarity 54.5%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
Db 266 EVILPGFDYS 276
:::|:|

RESULT 17
Q9YFI3
ID Q9YFI3 PRELIMINARY; PRT; 440 AA.
AC Q9YFI3
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 440AA long hypothetical alkaline protease.
GN APE0263.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=K1;
RX MEDLINE-99310339; PubMed-10382966;
RA Kavarabayasi Y.; Hino Y.; Horikawa H.; Yamazaki S.; Haikawa Y.;
RA Jin-no K.; Takahashi M.; Sekine M.; Baba S.-I.; Ankai A.; Kosugi H.;
RA Hosoyma A.; Fukui S.; Nagai Y.; Nishijima K.; Nakazawa H.;
RA Takamiya M.; Masuda S.; Funahashi T.; Tanaka T.; Kudoh Y.;
RA Yamazaki J.; Kishida N.; Oguchi A.; Aoki K.-I.; Kubota K.;
RA Nakamura Y.; Nomura N.; Sako Y.; Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000058; BAA79178.1;
DR HSSP: P00782; 2SET.
DR InterPro: IPR00209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 440 AA; 44680 MW; E43E6E2174B6F07E CRC64;

Query Match 64.8%; Score 35; DB 17; Length 440;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMDY 10
Db 120 EVLPWGVY 128
:::|:|

RESULT 18
O52680
ID O52680 PRELIMINARY; PRT; 511 AA.
AC O52680; O47289;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
GN RHSG.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR-50;
RA Wang Y.-D.; Zhao S.; Hill C.W.;

RT "Rhs elements comprise three subfamilies";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; AF044506; AAC62388.1; -;
 FT NON_TER 511 511
 SQ SEQUENCE 511 AA; 55877 MW; C85402569450DBFF CRC64;

Query Match 64.8%; Score 35; DB 2; Length 511;
 Best Local Similarity 60.0%; Pred. No. 75;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 DB 381 EQVNPGLDY 390
 I:|:|:|:|

RESULT 19
 Q8XZL5 PRELIMINARY; PRT; 517 AA.
 ID Q8XZL5;
 AC Q8XZL5;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Probable SUBSTATE-binding periplasmic (PBP) ABC transporter protein.
 DE RSC1380 OR RS04663.
 GN Ralstonia solanacearum (Pseudomonas solanacearum).
 OS Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S., Arlat M., Billault A., Brottier P., Canus J.C., Cattolico L., Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N., Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T., Signier P., Thebault P., Whalen M., Wincker P., Levy M., Weissenbach J., Boucher C.A.;
 RA "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL; AL640664; CAD15082.1; -;
 DR InterPro: IPR000914; SBP_bac_5.
 DR Pfam: PF00496; SBP_bac_5; 1.
 KW Complete proteome.
 SQ SEQUENCE 517 AA; 57349 MW; 861EE836F645F2DA CRC64;

Query Match 64.8%; Score 35; DB 16; Length 517;
 Best Local Similarity 66.7%; Pred. No. 76;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 DB 316 VVPQGVDA 324
 I:|:|:|:|

RESULT 20
 Q95P46 PRELIMINARY; PRT; 745 AA.
 ID Q95P46;
 AC Q95P46;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Na+/K+/2Cl-cotransporter (Fragment).
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-GILL;
 RA Weihrauch D., Towle D.W.;

RT "Sequence and expression analysis of the Na+/K+/2Cl- cotransporter in the euryhaline crabs Callinectes sapidus and Carcinus maenas";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; AY035548; AAK62044.1; -;
 DR InterPro: IPR002293; AA/rel_primeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR NON_TER 745 745
 FT NON_TER 1
 SQ SEQUENCE 745 AA; 81249 MW; 08AE2D3E1F70C7BE CRC64;

Query Match 64.8%; Score 35; DB 5; Length 745;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMDYS 11
 DB 625 VPOGLDYS 632
 I:|:|:|:|

RESULT 21
 Q9U6A3 PRELIMINARY; PRT; 1031 AA.
 ID Q9U6A3;
 AC Q9U6A3;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Na+/K+/2Cl-cotransporter.
 GN NKCC.
 OS Callinectes sapidus (Blue crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Callinectes.
 OX NCBI_TaxID=6763;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-GILL;
 RA Towle D.W.;
 RA "Cloning and sequencing a Na+/K+/2Cl- cotransporter from gills of the euryhaline blue crab Callinectes sapidus";
 RL Am. Zoologist 38:114A-114A(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-GILL;
 RA Towle D.W., Weihrauch D.;
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; AF190129; AAF05702.1; -;
 DR InterPro: IPR002293; AA/rel_primeasel.
 DR InterPro: IPR004842; KCL_cotransport.
 DR InterPro: IPR002443; NaKCL_transprter.
 DR PRINTS: PR01207; NAKCLTRNSPRT.
 DR TIGRFAMS: TIGR00930; 2a30; 1.
 SQ SEQUENCE 1031 AA; 113699 MW; F67773021AB05D71 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 1031;
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMDYS 11
 DB 737 VPOGLDYS 744
 I:|:|:|:|

RESULT 22
 Q17704 PRELIMINARY; PRT; 1150 AA.
 ID Q17704;
 AC Q17704; O17976;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE C53A5.2 protein.
 GN C53A5.2.
 OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mortimore B.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z81486; CAB03994.1; -.
 DR EMBL; Z78015; CAB03994.1; JOINED.
 DR EMBL; Z78015; CAB01437.1; -.
 DR EMBL; Z81486; CAB01437.1; JOINED.
 DR InterPro; IPR001201; PAP_25A_core.
 DR InterPro; IPR000051; SAM_bind.
 DR InterPro; IPR003402; Unk_Met10.
 DR Pfam; PF02475; Met_10; 1.
 SQ SEQUENCE 1150 AA; 133001 MW; BA130A251C1A12F8 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 1150;
 Best Local Similarity 66.7%; Pred. NO. 1.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVFXGMDYS 11
 :|||:
 DB 562 VLPVGIDYS 570

RESULT 23
 O52673
 ID 052673 PRELIMINARY; PRT; 1410 AA.
 AC 052673;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Core protein.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC11;
 RX MEDLINE=98361897; PubMed=9696756;
 RA Wang Y.D., Zhao S., Hill C.W.;
 RT "Rhs elements comprise three subfamilies which diverged prior to
 RT acquisition by Escherichia coli.";
 RL J. Bacteriol. 180:4102-4110(1998).
 DR EMBL; AF044503; AAC32478.1; -.
 DR InterPro; IPR001826; Rhs.
 DR Pfam; PF03527; Rhs; 1.
 DR PRINTS; PR00394; RHSPROTEIN.
 SQ SEQUENCE 1410 AA; 158739 MW; 39ECB9F8C97FCB15 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 1410;

Best Local Similarity 60.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 :|||:
 DB 382 EQVNPGLDY 391

RESULT 24
 O52666
 ID 052666 PRELIMINARY; PRT; 1420 AA.
 AC 052666;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Core protein.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC45;
 RX MEDLINE=98361897; PubMed=9696756;
 RA Wang Y.D., Zhao S., Hill C.W.;
 RT "Rhs elements comprise three subfamilies which diverged prior to
 RT acquisition by Escherichia coli.";
 RL J. Bacteriol. 180:4102-4110(1998).
 DR EMBL; AF044501; AAC32471.1; -.
 DR InterPro; IPR001826; Rhs.
 DR InterPro; IPR000130; Zn_MTPeptidse.
 DR Pfam; PF03527; Rhs; 1.
 DR PRINTS; PR00394; RHSPROTEIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1420 AA; 159215 MW; E7851E4D48740621 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 1420;
 Best Local Similarity 60.0%; Pred. NO. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 :|||:
 DB 381 EQVNPGLDY 390

RESULT 25
 O27146
 ID 027146 PRELIMINARY; PRT; 1474 AA.
 AC 027146;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative membrane protein.
 GN MTH1074.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 RA Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Iwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000878; AAB85563.1; -.

DR InterPro: IPR001434; DUF11.
DR Pfam: PF01345; DUF11; 9.
KW Complete proteome.
SQ SEQUENCE 1474 AA; 153713 MW; 11D4D27BD86255F0 CRC64;
Query Match 64.8%; Score 35; DB 17; Length 1474;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 EVVPXGMDY 10
Db 1238 DVLPAGLDY 1246

Search completed: June 10, 2003, 13:46:33
Job time : 25.7857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds
(without alignments)
46.744 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 EEVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	52	96.3	11	23	ABB80524
2	52	96.3	11	23	ABB80528
3	52	96.3	11	23	ABB80529
4	52	96.3	11	23	ABB80561
5	52	96.3	11	23	ABB80562
6	47	87.0	11	23	ABB80538
7	47	87.0	11	23	ABB80542
8	47	87.0	11	23	ABB80543
9	46	85.2	11	23	ABB80521
10	46	85.2	11	23	ABB80522

11	46	85.2	11	23	ABB80525	Hepatitis C virus
12	46	85.2	11	23	ABB80526	Hepatitis C virus
13	46	85.2	11	23	ABB80547	Hepatitis C virus
14	46	85.2	11	23	ABB80548	Hepatitis C virus
15	46	85.2	11	23	ABB80551	Hepatitis C virus
16	46	85.2	11	23	ABB80556	Hepatitis C virus
17	46	85.2	11	23	ABB80557	Hepatitis C virus
18	46	85.2	11	23	ABB80559	Hepatitis C virus
19	46	85.2	11	23	ABB80563	Hepatitis C virus
20	46	85.2	11	23	ABB80564	Hepatitis C virus
21	46	85.2	11	23	ABB80565	Hepatitis C virus
22	46	85.2	11	23	ABB80566	Hepatitis C virus
23	46	85.2	11	23	ABB80567	Hepatitis C virus
24	46	85.2	11	23	ABB80568	Hepatitis C virus
25	45	83.3	11	23	ABB80523	Hepatitis C virus
26	45	83.3	11	23	ABB80527	Hepatitis C virus
27	45	83.3	11	23	ABB80538	Hepatitis C virus
28	45	83.3	11	23	ABB80560	Hepatitis C virus
29	44	81.5	11	23	ABB80533	Hepatitis C virus
30	44	81.5	11	23	ABB80534	Hepatitis C virus
31	41	75.9	11	23	ABB80535	Hepatitis C virus
32	41	75.9	11	23	ABB80536	Hepatitis C virus
33	41	75.9	11	23	ABB80539	Hepatitis C virus
34	41	75.9	11	23	ABB80540	Hepatitis C virus
35	40	74.1	11	23	ABB80537	Hepatitis C virus
36	40	74.1	11	23	ABB80541	Hepatitis C virus
37	40	74.1	11	23	ABB80544	Hepatitis C virus
38	40	74.1	11	23	ABB80545	Hepatitis C virus
39	40	74.1	11	23	ABB80549	Hepatitis C virus
40	40	74.1	11	23	ABB80552	Hepatitis C virus
41	40	74.1	11	23	ABB80553	Hepatitis C virus
42	39	72.2	11	23	ABB80546	Hepatitis C virus
43	39	72.2	11	23	ABB80550	Hepatitis C virus
44	39	72.2	11	23	ABB80554	Hepatitis C virus
45	39	72.2	11	23	ABB80555	Hepatitis C virus
46	38	70.4	11	23	ABB80530	Hepatitis C virus
47	37	68.5	11	23	ABB80532	Hepatitis C virus
48	37	68.5	11	23	AAU76810	Citrus partial suc
49	36	66.7	20	20	AAU76810	Citrus partial suc
50	36	66.7	341	22	AAE00907	Herbicidally activ
51	36	66.7	348	22	AAE00908	Novel human diagno
52	36	66.7	460	23	ABG03621	Novel human diagno
53	36	66.7	1022	22	ABG03621	SFS protein. Zea
54	36	66.7	1022	22	ABG05826	Sucrose phosphate
55	36	66.7	1022	22	ABG08173	Maize sucrose phos
56	36	66.7	1068	13	AAU76810	Rice sucrose phosp
57	36	66.7	1068	13	AAU76810	Herbicidally activ
58	36	66.7	1068	18	AAW38266	Human polypeptide
59	36	66.7	1068	18	AAW38266	Mab L243 VH region
60	36	66.7	1071	15	AAW38266	Humanized antibody
61	36	66.7	1071	15	AAW38266	CDR-grafted L243-g
62	35	64.8	2778	22	ABG05826	MHC-II Mab L243 he
63	34.5	63.9	748	22	ABG05826	Drosophila melanog
64	34	63.0	140	16	AAU40225	Human polypeptide
65	34	63.0	140	16	AAU40225	Mab L243 VH region
66	34	63.0	140	16	AAU40225	Humanized antibody
67	34	63.0	140	16	AAU40225	CDR-grafted L243-g
68	34	63.0	143	22	ABG05826	MHC-II Mab L243 he
69	34	63.0	150	22	AAU40225	Drosophila melanog
70	34	63.0	150	22	AAU40225	Arabidopsis thalia
71	34	63.0	222	22	AAU40225	Bovine mammary tis
72	34	63.0	222	22	AAU40225	Group B Streptococ
73	34	63.0	565	23	ABP26468	Streptococcus poly
74	34	63.0	587	21	AAU74287	Lactococcus lactis
75	34	63.0	1062	22	AAU33615	Neisseria meningit

ALIGNMENTS

RESULT 1
ABB80524
ID ABB80524 standard; peptide; 11 AA.

FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8
 FT /note= "D-form residue"
 FT Misc-difference 9
 FT /note= "D-form residue"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 FT WO200208251-A2.
 XX
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 PR (CORV-) CORVAS INT INC.
 XX
 PA Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 96.3%; Score 52; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMDYS 11
 RESULT 4
 ABB80561
 ID ABB80561 standard; peptide; 11 AA.
 XX
 AC ABB80561;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT

FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 PR (CORV-) CORVAS INT INC.
 XX
 PA Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX
 PS Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 96.3%; Score 52; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMDYS 11
 RESULT 5
 ABB80562
 ID ABB80562 standard; peptide; 11 AA.
 XX
 AC ABB80562;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN

XX PD 31-JAN-2002.
 XX PF 19-JUL-2001; 2001WO-US23169.
 XX PR 21-JUL-2000; 2000US-220101P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX PS WPI; 2002-361643/39.
 XX DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX PT activity useful for treating disorders associated with hepatitis C
 XX PT virus protease -
 XX PS Claim 17; Page 65; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have
 XX CC virucide activity, and are useful for treating and in the manufacture of
 XX CC a medicament to treat disorders associated with HCV protease. A
 XX CC pharmaceutical composition comprising the peptide as an active ingredient
 XX CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 96.3%; Score 52; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMDYS 11
 Db 1 EEVVPXGMDYS 11
 RESULT 6
 ABB80538
 ID ABB80538 standard; peptide; 11 AA.
 AC ABB80538;
 AC DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Misc-difference 9 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX PD 19-JUL-2001; 2001WO-US23169.
 XX PF 21-JUL-2000; 2000US-220101P.
 XX PR (CORV-) CORVAS INT INC.
 XX PA

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX PT activity useful for treating disorders associated with hepatitis C
 XX PT virus protease -
 XX PS Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have
 XX CC virucide activity, and are useful for treating and in the manufacture of
 XX CC a medicament to treat disorders associated with HCV protease. A
 XX CC pharmaceutical composition comprising the peptide as an active ingredient
 XX CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 87.0%; Score 47; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMDYS 11
 Db 1 EEVVPXGMDYS 11
 RESULT 7
 ABB80542
 ID ABB80542 standard; peptide; 11 AA.
 AC ABB80542;
 AC DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Misc-difference 8 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX PD 19-JUL-2001; 2001WO-US23169.
 XX PF 21-JUL-2000; 2000US-220101P.
 XX PR (CORV-) CORVAS INT INC.
 XX PA Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX PT activity useful for treating disorders associated with hepatitis C
 XX PT virus protease -

OY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11

RESULT 10

ABB80522
 ID ABB80522 standard; peptide; 11 AA.

AC ABB80522;
 XX
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 residue 7"
 FT Misc-difference 9
 FT /note= "D-form residue"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11

RESULT 11

ABB80525
 ID ABB80525 standard; peptide; 11 AA.

AC ABB80525;
 XX
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 residue 7"
 FT Misc-difference 8
 FT /note= "D-form residue"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11

RESULT 12

ABB80526
 ID ABB80526 standard; peptide; 11 AA.

AC ABB80526;
 XX
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW	virucide.
XX	
XX	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "N-terminal acetyl"
FT	Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"
FT	Misc-difference 8 /note= "D-form residue"
FT	Misc-difference 9 /note= "D-form residue"
FT	Modified-site 11 /note= "C-terminal amide"
FT	
XX	WO200208251-A2.
PN	31-JAN-2002.
XX	
PD	19-JUL-2001; 2001WO-US23169.
PF	
XX	21-JUL-2000; 2000US-220101P.
PR	(CORV-) CORVAS INT INC.
XX	
PA	Lim-wilby M, Levy OE, Brunck TK;
PI	WPI; 2002-361643/39.
PT	Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease
PT	
XX	Claim 17; Page 64; 69pp; English.
XX	The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a pharmaceutical to treat disorders associated with HCV protease. A pharmacological composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.
XX	
SC	Sequence 11 AA;
Query Match	85.2%; Score 46; DB 23; Length 11;
Best Local Similarity	90.9%; pred. No. 0.022;
Matches 10; Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	1 EEVVPXGMDYS 11 11
Dd	1 EEVVPXGMSYS 11
RESULT 13	
ABB80547	
ID	ABB80547 standard; peptide; 11 AA.
XX	
AC	ABB80547;
XX	
DT	08-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #27.
XX	
KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW	virucide.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers

PN WO200208251-A2.
 XX 31-JAN-2002.
 PD 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease.
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB ||||| |||
 1 EEVVPXGTDYS 11

RESULT 15
 ABB80551
 ID ABB80551 standard; peptide; 11 AA.
 AC ABB80551;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.

PA (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease.
 XX Claim 17; Page 65; 69pp; English.
 PS The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB ||||| |||
 1 EEVVPXGTDYS 11

RESULT 16
 ABB80556
 ID ABB80556 standard; peptide; 11 AA.
 AC ABB80556;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 PA Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C

PT virus protease .
 XX Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGSDYS 11
 RESULT 17
 ABB80557
 ID ABB80557 standard; peptide; 11 AA.
 XX AC ABB80557;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease .
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.

CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGSDYS 11
 RESULT 18
 ABB80559
 ID ABB80559 standard; peptide; 11 AA.
 XX AC ABB80559;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease .
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 |||||
 Db 1 EEVVPXGMSYS 11

RESULT 19

ABB80563
 ID ABB80563 standard; peptide; 11 AA.

XX AC ABB80563;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1

FT FT /note= "N-terminal acetyl"

FT FT Modified-site 6

FT FT /note= "Valyl carbonyl forming keto-amide linkage with
 residue 7"

FT FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX DR WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 activity useful for treating disorders associated with hepatitis C
 virus protease -

XX PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 invention are alpha-ketoamide peptide analogues. The peptides have
 virucide activity, and are useful for treating and in the manufacture of
 a medicament to treat disorders associated with HCV protease. A
 pharmaceutical composition comprising the peptide as an active ingredient
 is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 |||||
 Db 1 EEVVPXGMSYS 11

RESULT 20

ABB80564
 ID ABB80564 standard; peptide; 11 AA.

XX AC ABB80564;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1

FT FT /note= "N-terminal acetyl"

FT FT Modified-site 6

FT FT /note= "Leucyl carbonyl forming keto-amide linkage with
 residue 7"

FT FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX DR WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 activity useful for treating disorders associated with hepatitis C
 virus protease -

XX PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 invention are alpha-ketoamide peptide analogues. The peptides have
 virucide activity, and are useful for treating and in the manufacture of
 a medicament to treat disorders associated with HCV protease. A
 pharmaceutical composition comprising the peptide as an active ingredient
 is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 |||||
 Db 1 EEVVPXGMSYS 11

RESULT 21

ABB80565
 ID ABB80565 standard; peptide; 11 AA.

XX AC ABB80565;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 XX Synthetic.
 XX
 XX OS
 XX Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminal acetyl"
 FT 6
 FT Modified-site
 FT /note= "Norleucyl carbonyl forming keto-amide linkage
 FT with residue 7"
 FT
 FT Modified-site 11
 FT /note= "C-terminal amide"
 FT
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 PF
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX Sequence 11 AA;
 SQ
 Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11
 XX
 XX 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminal acetyl"
 FT 6
 FT Modified-site
 FT /note= "Norleucyl carbonyl forming keto-amide linkage with residue 7"
 FT
 FT Modified-site 11
 FT /note= "C-terminal amide"
 FT
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 PF
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX Sequence 11 AA;
 SQ
 Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11
 XX
 XX 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminal acetyl"
 FT 6
 FT Modified-site

FT /note= "2-aminoisobutryl carbonyl residue forming a
 FT keto-amide linkage with residue 7"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 PF
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX Sequence 11 AA;
 SQ
 Query Match 85.2%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.022; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB 1 EEVVPXGMSYS 11
 XX
 XX 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminal acetyl"
 FT 6
 FT Modified-site
 FT /note= "(S,S)allothreonyl carbonyl residue forming a
 FT keto-amide linkage with residue 7"
 FT
 FT Modified-site 11
 FT /note= "C-terminal amide"
 FT
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 PF

XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX SQ Sequence 11 AA;
 XX
 XX Query Match 85.2%; Score 46; DB 23; Length 11;
 XX Best Local Similarity 90.9%; Pred. No. 0.022;
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB |||||
 1 EEVVPXGMSYS 11
 RESULT 24
 ABB80568
 ID ABB80568 standard; peptide; 11 AA.
 XX ABB80568;
 XX
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Alpha-propionyl-glycyl-carbonyl residue forming
 FT a keto-amide linkage with residue 7"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX SQ Sequence 11 AA;
 XX
 XX Query Match 85.2%; Score 46; DB 23; Length 11;
 XX Best Local Similarity 90.9%; Pred. No. 0.022;
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 DB |||||
 1 EEVVPXGMSYS 11
 RESULT 25
 ABB80523
 ID ABB80523 standard; peptide; 11 AA.
 XX ABB80523;
 XX
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX

SQ Sequence 11 AA;

Query Match 83.3%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.034;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||||

Db 1 EEVVPXGMHYS 11

Search completed: June 10, 2003, 13:39:09

Job time : 31.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds
(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 EHVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	66.7	341	US-08-853-948B-4	Sequence 4, Appli
2	36	66.7	348	US-08-853-948B-5	Sequence 5, Appli
3	36	66.7	368	US-09-697-367-24	Sequence 21, Appli
4	36	66.7	1068	US-08-429-054A-11	Sequence 11, Appli
5	36	66.7	1068	US-08-718-777-7	Sequence 7, Appli
6	36	66.7	1068	US-09-051-341-7	Sequence 7, Appli
7	34	63.0	140	US-08-569-147-76	Sequence 76, Appli
8	34	63.0	140	US-08-569-147-82	Sequence 82, Appli
9	33	61.1	59	US-08-963-851-14	Sequence 14, Appli
10	33	61.1	378	US-08-070-165F-8	Sequence 8, Appli
11	33	61.1	378	US-08-885-418-8	Sequence 8, Appli
12	33	61.1	801	US-09-383-630-6	Sequence 6, Appli
13	32	59.3	65	5177197-51	Patent No. 5177197
14	32	59.3	102	US-08-580-988A-23	Sequence 23, Appli
15	32	59.3	152	US-08-460-694-4	Sequence 4, Appli
16	32	59.3	152	US-08-460-744-4	Sequence 4, Appli
17	32	59.3	152	US-07-667-711B-4	Sequence 4, Appli
18	32	59.3	173	US-08-193-977-7	Sequence 7, Appli
19	32	59.3	189	US-08-464-517-21	Sequence 21, Appli
20	32	59.3	189	US-08-246-361A-21	Sequence 21, Appli
21	32	59.3	189	US-08-463-772-21	Sequence 21, Appli
22	32	59.3	189	PCT-US93-05000-21	Sequence 21, Appli
23	32	59.3	231	US-08-926-842B-20	Sequence 20, Appli
24	32	59.3	236	US-08-464-517-22	Sequence 22, Appli
25	32	59.3	236	US-08-246-361A-22	Sequence 22, Appli
26	32	59.3	236	US-08-463-772-22	Sequence 22, Appli
27	32	59.3	236	PCT-US93-05000-22	Sequence 22, Appli

Sequence 21, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 23, Appli
Sequence 23, Appli
Sequence 23, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 19, Appli
Sequence 19, Appli
Sequence 20, Appli
Sequence 20, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 19, Appli
Sequence 20, Appli
Sequence 4, Appli
Patent No. 5177197
Sequence 10, Appli
Sequence 3, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 4, Appli
Sequence 5, Appli
Patent No. 5177197
Sequence 5124, Ap
Sequence 57, Appli
Sequence 3003, Ap
Sequence 4, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-853-948B-4
; Sequence 4, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853.948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Citrus unshiu
; FEATURE:
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asp, Asn, Phe,
; Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe,

OTHER INFORMATION: Pro, Ser, Thr, Tyr, or Val
US-08-853-948B-4

Query Match 66.7%; Score 36; DB 4; Length 341;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |:|:|
Db 228 VIPPGMDFS 236

RESULT 2

US-08-853-948B-5
; Sequence 5, Application US/08853948B
; Patent No. 6210943

GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING

; FILE REFERENCE: 0049-0235-0

; CURRENT APPLICATION NUMBER: US/08/853,948B

; CURRENT FILING DATE: 1997-05-09

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 5

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Citrus unshiu

US-08-853-948B-5

Query Match

Best Local Similarity 66.7%; Score 36; DB 4; Length 348;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11

|:| |:|:|
Db 234 VIPPGMDFS 242

RESULT 3

US-09-697-367-24

; Sequence 24, Application US/09697367

; Patent No. 6323015

GENERAL INFORMATION:

; APPLICANT: Orozco Jr., Emil M.

; APPLICANT: Calmi, Perry G.

; APPLICANT: Weng, Zude

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE

; FILE REFERENCE: BB1166 US NA

; CURRENT APPLICATION NUMBER: US/09/697,367

; CURRENT FILING DATE: 2000-10-26

; PRIOR FILING DATE: 60/084,529

; PRIOR FILING DATE: 1998-MAY-07

; PRIOR APPLICATION NUMBER: PCT/US99/09865

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; LENGTH: 368

; TYPE: PRT

; ORGANISM: Zea mays

US-09-697-367-24

Query Match

Best Local Similarity 66.7%; Score 36; DB 4; Length 368;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11

|:| |:|:|
Db 217 VIPPGMDFS 225

RESULT 4

US-08-429-054A-11

; Sequence 11, Application US/08429054A

; Patent No. 5917126

GENERAL INFORMATION:

; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,

; APPLICANT: JEAN; VOELKER, TONI; GERVAIS, MONICA

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),

; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN AND MUSERLIAN

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/429,054A

; FILING DATE: 26-APR-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 842,337

; FILING DATE: 20-March-1992

; APPLICATION NUMBER: PCT/FR 91/00593

; FILING DATE: 18-July-1991

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: French 90402094.9

; FILING DATE: 20-July-1990

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Charles A. Muserlian

; REGISTRATION NUMBER: 19,683

; REFERENCE/DOCKET NUMBER: 146.1137

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000

; TELEFAX: (212) 661-8002

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1068

; TYPE: Amino acid

; STRANDEDNESS: Single

; TOPOLOGY: Unknown

; MOLECULE TYPE: Peptide

US-08-429-054A-11

Query Match 66.7%; Score 36; DB 2; Length 1068;

Best Local Similarity 66.7%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY

3 VVPXGMDYS 11

|:| |:|:|
Db 435 VIPPGMDFS 443

RESULT 5

US-08-718-777-7

; Sequence 7, Application US/08718777

; Patent No. 5981852

GENERAL INFORMATION:

; APPLICANT: Van Assche, C.

; APPLICANT: Lando, D.

; APPLICANT: Bruneau, J. M.

; APPLICANT: Voelker, T.

; APPLICANT: Gervais, M.

; TITLE OF INVENTION: MODIFICATION OF SUCROSE
 ; TITLE OF INVENTION: PHOSPHATE
 ; TITLE OF INVENTION: SYNTHASE IN PLANTS
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Offices of Barbara Rae-Venter
 ; STREET: 260 Sheridan Avenue, Suite 440
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/718,777
 ; FILING DATE: NOT YET ASSIGNED
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/175,471
 ; FILING DATE: 27-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barbara Rae-Venter
 ; REGISTRATION NUMBER: 32,750
 ; REFERENCE/DOCKET NUMBER: CGNE.072.02US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415)328-4400
 ; TELEFAX: (415)328-4477
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1068 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-718-777-7

Query Match 66.7%; Score 36; DB 2; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGMDYS 11
 I:|:||||:|
 DB 435 VVPGMDFS 443

RESULT 6
 US-09-051-341-7
 ; Sequence 7, Application US/09051341
 ; Patent No. 6124528
 ; GENERAL INFORMATION:
 ; APPLICANT: Shewmaker, C. K.
 ; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
 ; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rae-Venter Law Group, P.C.
 ; STREET: 260 Sheridan Avenue, Suite 440
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/051,341
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/17351
 ; FILING DATE: 25-OCT-1996
 ; APPLICATION NUMBER: US 08/549,016
 ; FILING DATE: 27-OCT-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/372,200
 ; FILING DATE: 12-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barbara Rae-Venter, Ph.D.,
 ; REGISTRATION NUMBER: 32,750
 ; REFERENCE/DOCKET NUMBER: CGNE.110.02US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415)328-4400
 ; TELEFAX: (415)328-4477
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1068 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-051-341-7

Query Match 66.7%; Score 36; DB 3; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGMDYS 11
 I:|:||||:|
 DB 435 VVPGMDFS 443

RESULT 7
 US-08-569-147-76
 ; Sequence 76, Application US/08569147
 ; Patent No. 6180377
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: HUMANISED ANTIBODIES
 ; NUMBER OF SEQUENCES: 95
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
 ; ADDRESS: No. 6180377ris, LLP
 ; STREET: One Liberty Place - 46th Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/569,147
 ; FILING DATE: 25-March-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Trujillo, Doreen Yanko
 ; REGISTRATION NUMBER: 35,719
 ; REFERENCE/DOCKET NUMBER: CARP-0047
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 76:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 140 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-569-147-76

Query Match 63.0%; Score 34; DB 4; Length 140;
 Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDY 10
| | | | |
Db 122 VVPTGFDY 129

RESULT 8

US-08-569-147-82
; Sequence 82, Application US/08569147
; Patent No. 6180377

; GENERAL INFORMATION:

; APPLICANT: HUMANISED ANTIBODIES

; TITLE OF INVENTION: HUMANISED ANTIBODIES

; NUMBER OF SEQUENCES: 95

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: No. 6180377ris, LLP

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: U.S.A.

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/569,147

; FILING DATE: 25-March-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Trujillo, Doreen Yanko

; REGISTRATION NUMBER: 35,719

; REFERENCE/DOCKET NUMBER: CARP-0047

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 82:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 140 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-569-147-82

Query Match 63.0%; Score 34; DB 4; Length 140;

Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDY 10

| | | | |

Db 122 VVPTGFDY 129

RESULT 9

US-08-963-851-14

; Sequence 14, Application US/08963851

; Patent No. 6300116

; GENERAL INFORMATION:

; APPLICANT: VAN DER OSTEN, CLAUD

; APPLICANT: HALKIER, TORDEN

; APPLICANT: ANDERSEN, CARSTEN

; APPLICANT: BAUDITZ, PETER

; APPLICANT: HANSEN, PETER KAMP

; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS

; FILE REFERENCE: 4946,200-US

; CURRENT APPLICATION NUMBER: US/08/963,851

; CURRENT FILING DATE: 1997-11-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 14

; LENGTH: 59

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-08-963-851-14

Query Match 61.1%; Score 33; DB 4; Length 59;

Best Local Similarity 45.5%; Pred. No. 7.5;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

| | | | |

Db 38 EXHIPGGLYS 48

RESULT 10

US-08-070-165F-8

; Sequence 8, Application US/08070165F

; Patent No. 5750365

; GENERAL INFORMATION:

; APPLICANT: Chiu, Ing-Ming

; APPLICANT: Poulin, Matthew L

; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ing-Ming Chiu

; STREET: S2052 Davis Medical Research Center, 480 West

; CITY: Columbus

; STATE: Ohio

; COUNTRY: USA

; ZIP: 43210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/070,165F

; FILING DATE:

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (614)-293-8093

; TELEFAX: (614)-293-5631

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 378 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-070-165F-8

Query Match 61.1%; Score 33; DB 1; Length 378;

Best Local Similarity 85.7%; Pred. No. 64;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11

| | | | |

Db 139 PPGMDYS 145

RESULT 11

US-08-885-418-8

; Sequence 8, Application US/08885418

; Patent No. 5925528

; GENERAL INFORMATION:

; APPLICANT: Chiu, Ing-Ming

; APPLICANT: Poulin, Matthew L

; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ing-Ming Chiu

; STREET: S2052 Davis Medical Research Center, 480 West

; CITY: Columbus

; STATE: Ohio

; COUNTRY: USA

; ZIP: 43210

```
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-885-418-8

Query Match 61.1%; Score 33; DB 2; Length 378;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
| | | | |
Db 139 PPGMDYS 145

RESULT 12
US-09-383-630-6
; Sequence 6, Application US/09383630A
; Patent No. 6265632
; GENERAL INFORMATION:
; APPLICANT: Avner Yayon et al.
; TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH
; FACTOR RECEPTOR ASSOCIATED
; CHONDRODYSPLASIA
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted
; to an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,630A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 1402/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-562553
; TELEFAX: 972-3-562554
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 801
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-383-630-6

Query Match 61.1%; Score 33; DB 4; Length 801;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
| | | | |
Db 566 PPGMDYS 572

RESULT 13
5177197-51
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,
; LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 51:
; LENGTH: 65
5177197-51

Query Match 59.3%; Score 32; DB 6; Length 65;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
| | | | |
Db 52 KEICPGMGYT 62

RESULT 14
US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5856161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; RECEPTOR-I-Associated Protein Kinase And Methods
; TITLE OF INVENTION: For Its Use
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
```

REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5721CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-777-2321
TELEFAX: 713-777-6908
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: no
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-580-988A-23

Query Match 59.3%; Score 32; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
DB 24 EEVFPPLAMNY 33

RESULT 15
US-08-460-694-4
; Sequence 4, Application US/08460694
; Patent No. 5858655
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: Pradl Cyclin and its cdna
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,694
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

Query Match 59.3%; Score 32; DB 2; Length 152;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10

Db 20 EEVFPPLAMNY 29
RESULT 16
US-08-460-744-4
; Sequence 4, Application US/08460744
; Patent No. 6107541
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: Pradl Cyclin and its cdna
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,744
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

Query Match 59.3%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
DB 20 EEVFPPLAMNY 29

RESULT 17
US-07-667-711B-4
; Sequence 4, Application US/07667711B
; Patent No. 6110700
; GENERAL INFORMATION:
; APPLICANT: ARNOLD, ANDREW
; TITLE OF INVENTION: Pradl Cyclin and its cdna
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/667,711B
;; FILING DATE: 11-MAR-1991
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MCPHAIL, DONALD R.
;; REGISTRATION NUMBER: 35,811
;; REFERENCE/DOCKET NUMBER: 0609.4070000
;; TELEPHONE: (202) 371-2600
;; TELEFAX: (202) 371-2540
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 152 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-07-667-711B-4

Query Match 59.3%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
||| | | |
DB 20 EEVFPLAMNY 29

RESULT 18
US-08-193-977-7
;; Sequence 7, Application US/08193977
;; Patent No. 5625031
;; GENERAL INFORMATION:
;; APPLICANT: WEBSTER, KEVIN R.
;; APPLICANT: COLEMAN, KEVIN G.
;; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND
;; TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN
;; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN
;; NUMBER OF SEQUENCES: 34
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: REED & ROBINS
;; STREET: 635 BRYANT STREET
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: UNITED STATES OF AMERICA
;; ZIP: 94301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/193,977
;; FILING DATE: 08-FEB-1994
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ROBINS, ROBERTA L.
;; REGISTRATION NUMBER: 33,208
;; REFERENCE/DOCKET NUMBER: 5998-0016
;; TELEPHONE: (415) 617-8999
;; TELEFAX: (415) 327-3231
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 173 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-193-977-7
Query Match 59.3%; Score 32; DB 1; Length 173;

Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
||| | | |
DB 55 EEVFPLAMNY 64

RESULT 19
US-08-464-517-21
;; Sequence 21, Application US/08464517
;; Patent No. 5869640
;; GENERAL INFORMATION:
;; APPLICANT: BEACH, DAVID H.
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII(text)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/464,517
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/963,308
;; FILING DATE: 16-OCT-1992
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-464-517-21

Query Match 59.3%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
||| | | |
DB 74 EEVFPLAMNY 83

RESULT 20
US-08-246-361A-21
;; Sequence 21, Application US/08246361A
;; Patent No. 5998582
;; GENERAL INFORMATION:
;; APPLICANT: BEACH, DAVID H.
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
;; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/246,361A
 ; FILING DATE: 19-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; APPLICATION NUMBER: US 07/701,514
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MII-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 189 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-246-361A-21

Query Match 59.3%; Score 32; DB 2; Length 189;
 Best Local Similarity 60.0%; Pred. No. 46;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 Db 74 EEVFLPMNY 83

RESULT 21
 US-08-463-772-21
 ; Sequence 21, Application US/08463772
 ; Patent No. 6066501
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,772
 ; FILING DATE:

; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/701,514
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MII-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 189 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-463-772-21

Query Match 59.3%; Score 32; DB 3; Length 189;
 Best Local Similarity 60.0%; Pred. No. 46;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
 Db 74 EEVFLPMNY 83

RESULT 22
 PCT-US93-05000-21
 ; Sequence 21, Application PC/TUS9305000
 ; GENERAL INFORMATION:
 ; APPLICANT: MITOTIX
 ; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: US
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/05000
 ; FILING DATE: 19930525
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/888,178
 ; FILING DATE: 26-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Granahan, Patricia
 ; REGISTRATION NUMBER: 32,227
 ; REFERENCE/DOCKET NUMBER: CSHL91-02A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-861-6240
 ; TELEFAX: 616-861-9540
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 189 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein

PCT-US93-05000-21

Query Match 59.3%; Score 32; DB 5; Length 189;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
||| | |:
Db 74 EEVFLAMNY 83

RESULT 23

US-08-926-842B-20
; Sequence 20, Application US/08926842B
; Patent No. 6030807
; GENERAL INFORMATION:
; APPLICANT: Sa-No. 6030807ueira, Isabel
; APPLICANT: de Lencastre, Herminia
; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/926,842B
; FILING DATE: 10-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-089 N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
US-08-926-842B-20

Query Match 59.3%; Score 32; DB 3; Length 231;
Best Local Similarity 55.6%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
: | |:
Db 41 IKPSGVDYS 49

RESULT 24

US-08-464-517-22
; Sequence 22, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIT-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-22

Query Match 59.3%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
||| | |:
Db 20 EEVFLAMNY 29

RESULT 25

US-08-246-361A-22
; Sequence 22, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-246-361A-22

Query Match 59.3%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMDY 10
Db 20 EEVFFLAMNY 29

Search completed: June 10, 2003, 13:51:34
Job time : 9.64286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds
(without alignments)
75.710 Million cell updates/sec

Title: US-09-909-164-13
Perfect score: 54
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	66.7	1049	US-10-217-700-10	Sequence 10, Appl
2	36	66.7	1068	US-10-217-700-8	Sequence 8, Appl
3	36	66.7	1081	US-10-217-700-4	Sequence 4, Appl
4	36	66.7	1083	US-10-217-700-11	Sequence 11, Appl
5	36	66.7	1084	US-10-217-700-9	Sequence 9, Appl
6	35	64.8	440	US-09-813-408-27	Sequence 27, Appl
7	34	63.0	1062	US-09-815-242-5111	Sequence 5111, Ap
8	34	63.0	3472	US-10-027-806-4	Sequence 4, Appl
9	34	63.0	3472	US-10-034-623-4	Sequence 4, Appl
10	34	63.0	3472	US-10-027-801-4	Sequence 14, Appl
11	33	61.1	59	US-09-948-080-14	Sequence 4881, Ap
12	33	61.1	283	US-09-738-626-4881	Sequence 10697, A
13	33	61.1	299	US-09-815-242-10697	Sequence 526, App
14	33	61.1	736	US-09-978-295A-526	Sequence 526, App
15	33	61.1	736	US-09-978-697-526	Sequence 526, App
16	33	61.1	736	US-09-978-192A-526	Sequence 526, App
17	33	61.1	736	US-09-999-832A-526	Sequence 526, App
18	33	61.1	736	US-09-978-189-526	Sequence 526, App
19	33	61.1	736	US-10-174-590-420	Sequence 420, App

20	33	61.1	736	9	US-10-176-758-420	Sequence 420, App
21	33	61.1	736	9	US-10-175-737-420	Sequence 420, App
22	33	61.1	736	9	US-10-173-706-420	Sequence 420, App
23	33	61.1	736	9	US-10-175-738-420	Sequence 420, App
24	33	61.1	736	9	US-10-175-752-420	Sequence 420, App
25	33	61.1	736	9	US-10-176-482-420	Sequence 420, App
26	33	61.1	736	9	US-10-176-757-420	Sequence 420, App
27	33	61.1	736	9	US-10-176-913-420	Sequence 420, App
28	33	61.1	736	9	US-10-180-552-420	Sequence 420, App
29	33	61.1	736	9	US-10-180-557-420	Sequence 420, App
30	33	61.1	736	9	US-10-173-700-420	Sequence 420, App
31	33	61.1	736	9	US-10-174-579-420	Sequence 420, App
32	33	61.1	736	9	US-10-174-582-420	Sequence 420, App
33	33	61.1	736	9	US-10-174-588-420	Sequence 420, App
34	33	61.1	736	9	US-10-175-739-420	Sequence 420, App
35	33	61.1	736	9	US-10-175-740-420	Sequence 420, App
36	33	61.1	736	9	US-10-175-743-420	Sequence 420, App
37	33	61.1	736	9	US-10-176-488-420	Sequence 420, App
38	33	61.1	736	9	US-10-176-492-420	Sequence 420, App
39	33	61.1	736	9	US-10-176-747-420	Sequence 420, App
40	33	61.1	736	9	US-10-176-750-420	Sequence 420, App
41	33	61.1	736	9	US-10-176-985-420	Sequence 420, App
42	33	61.1	736	9	US-10-176-987-420	Sequence 420, App
43	33	61.1	736	9	US-10-176-991-420	Sequence 420, App
44	33	61.1	736	9	US-10-176-992-420	Sequence 420, App
45	33	61.1	736	9	US-10-176-993-420	Sequence 420, App
46	33	61.1	736	9	US-10-184-658-420	Sequence 420, App
47	33	61.1	736	9	US-10-173-695-420	Sequence 420, App
48	33	61.1	736	9	US-10-173-697-420	Sequence 420, App
49	33	61.1	736	9	US-10-173-705-420	Sequence 420, App
50	33	61.1	736	9	US-10-174-576-420	Sequence 420, App
51	33	61.1	736	9	US-10-174-585-420	Sequence 420, App
52	33	61.1	736	9	US-10-174-586-420	Sequence 420, App
53	33	61.1	736	9	US-10-175-747-420	Sequence 420, App
54	33	61.1	736	9	US-10-176-481-420	Sequence 420, App
55	33	61.1	736	9	US-10-176-485-420	Sequence 420, App
56	33	61.1	736	9	US-10-176-487-420	Sequence 420, App
57	33	61.1	736	9	US-10-176-493-420	Sequence 420, App
58	33	61.1	736	9	US-10-176-756-420	Sequence 420, App
59	33	61.1	736	9	US-10-176-911-420	Sequence 420, App
60	33	61.1	736	9	US-10-176-919-420	Sequence 420, App
61	33	61.1	736	9	US-10-176-925-420	Sequence 420, App
62	33	61.1	736	9	US-10-176-978-420	Sequence 420, App
63	33	61.1	736	9	US-10-179-510-420	Sequence 420, App
64	33	61.1	736	9	US-10-180-543-420	Sequence 420, App
65	33	61.1	736	9	US-10-180-544-420	Sequence 420, App
66	33	61.1	736	9	US-10-180-546-420	Sequence 420, App
67	33	61.1	736	9	US-10-180-547-420	Sequence 420, App
68	33	61.1	736	9	US-10-180-549-420	Sequence 420, App
69	33	61.1	736	9	US-10-180-555-420	Sequence 420, App
70	33	61.1	736	9	US-10-180-559-420	Sequence 420, App
71	33	61.1	736	9	US-10-181-000-420	Sequence 420, App
72	33	61.1	736	9	US-10-183-010-420	Sequence 420, App
73	33	61.1	736	9	US-10-183-012-420	Sequence 420, App
74	33	61.1	736	9	US-10-184-614-420	Sequence 420, App
75	33	61.1	736	9	US-10-184-614-420	Sequence 420, App

ALIGNMENTS

RESULT 1
US-10-217-700-10
; Sequence 10, Application US10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Holagler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12

; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-217-700-10

Query Match 66.7%; Score 36; DB 9; Length 1049;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||:|
Db 436 VIPPGMDFS 444

RESULT 2
US-10-217-700-8
; Sequence 8, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; EARLIER FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1068
; TYPE: PRT
; ORGANISM: Zea mays
US-10-217-700-8

Query Match 66.7%; Score 36; DB 9; Length 1068;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||:|
Db 435 VIPPGMDFS 443

RESULT 3
US-10-217-700-4
; Sequence 4, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Craterostigma plantagineum
US-10-217-700-4

Query Match 66.7%; Score 36; DB 9; Length 1081;

Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMDYS 11
|:| |||:|
Db 445 VIPPGMDFS 453

RESULT 4
US-10-217-700-11
; Sequence 11, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-217-700-11

Query Match 66.7%; Score 36; DB 9; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||:|
Db 483 VIPPGMDFS 491

RESULT 5
US-10-217-700-9
; Sequence 9, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-217-700-9

Query Match 66.7%; Score 36; DB 9; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:| |||:|
Db 453 VIPPGMDFS 461

RESULT 6
US-09-813-408-27
; Sequence 27, Application US/09813408

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds
(without alignments)
46.744 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EREVPGXSYS 11

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	45	93.8	11	23	ABB80544	Hepatitis C virus
2	45	93.8	11	23	ABB80545	Hepatitis C virus
3	45	93.8	11	23	ABB80549	Hepatitis C virus
4	45	93.8	11	23	ABB80552	Hepatitis C virus
5	45	93.8	11	23	ABB80553	Hepatitis C virus
6	44	91.7	11	23	ABB80521	Hepatitis C virus
7	44	91.7	11	23	ABB80522	Hepatitis C virus
8	44	91.7	11	23	ABB80525	Hepatitis C virus
9	44	91.7	11	23	ABB80526	Hepatitis C virus
10	44	91.7	11	23	ABB80530	Hepatitis C virus

11	44	91.7	11	23	ABB80535	Hepatitis C virus
12	44	91.7	11	23	ABB80536	Hepatitis C virus
13	44	91.7	11	23	ABB80539	Hepatitis C virus
14	44	91.7	11	23	ABB80540	Hepatitis C virus
15	44	91.7	11	23	ABB80559	Hepatitis C virus
16	44	91.7	11	23	ABB80563	Hepatitis C virus
17	44	91.7	11	23	ABB80564	Hepatitis C virus
18	44	91.7	11	23	ABB80565	Hepatitis C virus
19	44	91.7	11	23	ABB80566	Hepatitis C virus
20	44	91.7	11	23	ABB80567	Hepatitis C virus
21	44	91.7	11	23	ABB80568	Hepatitis C virus
22	41	85.4	11	23	ABB80547	Hepatitis C virus
23	41	85.4	11	23	ABB80548	Hepatitis C virus
24	41	85.4	11	23	ABB80551	Hepatitis C virus
25	41	85.4	11	23	ABB80556	Hepatitis C virus
26	41	85.4	11	23	ABB80557	Hepatitis C virus
27	40	83.3	11	23	ABB80524	Hepatitis C virus
28	40	83.3	11	23	ABB80528	Hepatitis C virus
29	40	83.3	11	23	ABB80529	Hepatitis C virus
30	40	83.3	11	23	ABB80533	Hepatitis C virus
31	40	83.3	11	23	ABB80534	Hepatitis C virus
32	40	83.3	11	23	ABB80538	Hepatitis C virus
33	40	83.3	11	23	ABB80542	Hepatitis C virus
34	40	83.3	11	23	ABB80543	Hepatitis C virus
35	40	83.3	11	23	ABB80546	Hepatitis C virus
36	40	83.3	11	23	ABB80550	Hepatitis C virus
37	40	83.3	11	23	ABB80554	Hepatitis C virus
38	40	83.3	11	23	ABB80555	Hepatitis C virus
39	40	83.3	11	23	ABB80561	Hepatitis C virus
40	40	83.3	11	23	ABB80562	Hepatitis C virus
41	39	81.2	11	23	ABB80523	Hepatitis C virus
42	39	81.2	11	23	ABB80527	Hepatitis C virus
43	39	81.2	11	23	ABB80531	Hepatitis C virus
44	39	81.2	11	23	ABB80532	Hepatitis C virus
45	39	81.2	11	23	ABB80537	Hepatitis C virus
46	39	81.2	11	23	ABB80541	Hepatitis C virus
47	39	81.2	11	23	ABB80558	Hepatitis C virus
48	39	81.2	11	23	ABB80560	Hepatitis C virus
49	38	79.2	1022	22	ABG03621	Novel human diagno
50	38	79.2	1022	22	ABG05826	Novel human diagno
51	38	79.2	1022	22	ABG08173	Novel human diagno
52	37	77.1	244	21	AA12881	Murine JNK3 bindin
53	37	77.1	484	21	AA12882	Murine JNK3 bindin
54	36	75.0	1070	22	AAU14378	Human novel protei
55	36	75.0	3472	21	AAV90913	Cenarchaeum symbio
56	34	70.8	20	20	AAU76810	Hepatitis C virus
57	34	70.8	398	20	AAV59885	Human normal uteru
58	34	70.8	440	21	AAE27201	Hela cell library
59	34	70.8	478	20	AAV02281	Secreted protein c
60	34	70.8	478	22	AAV39754	Human polypeptide
61	34	70.8	483	22	AAV41540	Human polypeptide
62	33	68.8	280	22	ABE68883	Drosophila melanog
63	33	68.8	299	22	ABG00168	Novel human diagno
64	32	66.7	109	23	ABG06850	Human nPCR-Seq106
65	32	66.7	410	22	AAV40649	Human polypeptide
66	32	66.7	413	21	AAV41846	Human ORFX ORF1610
67	32	66.7	602	23	ABV23116	Herbicidally activ
68	32	66.7	947	21	AAV25105	Pinus radiata cell
69	32	66.7	1732	23	ABV90754	Human Tumour Endot
70	32	66.7	4472	17	AAV97245	Virulence gene clu
71	31	64.6	25	23	ABG62372	Eubacterial DNA po
72	31	64.6	80	22	AAV92865	Human digestive sy
73	31	64.6	80	23	AAV52683	Human nucleotide r
74	31	64.6	99	22	AAV676703	Human colon cancer
75	31	64.6	159	18	AAV44125	Streptococcus pneu

ALIGNMENTS

RESULT 1
ABB80544
ID ABB80544 standard; peptide; 11 AA.

FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
PR 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
SQ
Query Match 93.8%; Score 45; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXSYS 11
DB 1 EEVVPXGSSYS 11
RESULT 4
ABB80552
ID ABB80552 standard; peptide; 11 AA.
XX
XX ABB80552;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #32.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
PA

XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
SQ
Query Match 93.8%; Score 45; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXSYS 11
DB 1 EEVVPXGSSYS 11
RESULT 5
ABB80553
ID ABB80553 standard; peptide; 11 AA.
XX
XX ABB80553;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #33.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
PA


```
ID ABB80530 standard; peptide; 11 AA.
XX
AC ABB80530;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #10.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX
OS virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 6 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 11 residue 7"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX
XX Claim 17; Page 64; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
XX Query Match 91.7%; Score 44; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.015;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EEVVPXGXSYS 11
XX |||||
XX Db 1 EEVVPXGQSYS 11
XX
XX RESULT 12
XX ABB80536
XX ID ABB80536 standard; peptide; 11 AA.
XX
XX AC ABB80536;
XX
XX DT 08-OCT-2002 (first entry)
XX
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #16.
XX
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX
XX OS virucide.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Misc-difference 9 residue 7"
XX
```

```
ID ABB80535 standard; peptide; 11 AA.
XX
AC ABB80535;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #15.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX
OS virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 6 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 11 residue 7"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX
XX Claim 17; Page 64; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
XX Query Match 91.7%; Score 44; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.015;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EEVVPXGXSYS 11
XX |||||
XX Db 1 EEVVPXGQSYS 11
XX
XX RESULT 11
XX ABB80535
XX ID ABB80535 standard; peptide; 11 AA.
XX
XX AC ABB80535;
XX
XX DT 08-OCT-2002 (first entry)
XX
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #15.
XX
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX
XX OS virucide.
```

FT Modified-site 11 /note= "D-form residue"
 FT 11 /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 PA
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX Claim 17; Page 64; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 |||||
 RESULT 13
 ABB80539
 ID ABB80539 standard; peptide; 11 AA.
 XX
 AC ABB80539;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 PA
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX Claim 17; Page 64; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 |||||
 RESULT 14
 ABB80540
 ID ABB80540 standard; peptide; 11 AA.
 XX
 AC ABB80540;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 PA
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 |||||

PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 |||||
 RESULT 14
 ABB80540
 ID ABB80540 standard; peptide; 11 AA.
 XX
 AC ABB80540;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 PA
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI
 XX WPI; 2002-361643/39.
 DR
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PT
 XX Claim 17; Page 65; 69pp; English.
 PS
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 |||||

PI Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 PS Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 ||||| |||
 RESULT 15
 ABB80559
 ID ABB80559 standard; peptide; 11 AA.
 AC ABB80559;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C

PT virus protease -
 XX
 CC Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 91.7%; Score 44; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 DB 1 EEVVPXGXSYS 11
 ||||| |||
 RESULT 16
 ABB80563
 ID ABB80563 standard; peptide; 11 AA.
 AC ABB80563;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 PS Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 91.7%; Score 44; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.015; Length 11;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGXSYS 11

RESULT 17

ABB80564

ID ABB80564 standard; peptide; 11 AA.

XX

AC ABB80564;

XX 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX Synthetic.

XX

PH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Leucyl carbonyl forming keto-amide linkage with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

PF 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

PR (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 91.7%; Score 44; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.015; Length 11;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGXSYS 11

RESULT 19

ABB80566

ID ABB80566 standard; peptide; 11 AA.

XX

AC ABB80566;

XX 08-OCT-2002 (first entry)

DB 1 EEVVPXGXSYS 11

RESULT 18

ABB80565

ID ABB80565 standard; peptide; 11 AA.

XX

AC ABB80565;

XX 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX Synthetic.

XX

PH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norleucyl carbonyl forming keto-amide linkage

FT with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

PF 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

PR (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 91.7%; Score 44; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.015;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGXSYS 11

RESULT 19

ABB80566

ID ABB80566 standard; peptide; 11 AA.

XX

AC ABB80566;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 6 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "2-aminoisobutyl carbonyl residue forming a
FT keto-amide linkage with residue 7"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 91.7%; Score 44; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXSYS 11
Db |||||
1 EEVVPXGMSYS 11
RESULT 20
ABB80567
ID ABB80567 standard; peptide; 11 AA.
XX
AC ABB80567;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1

FT Modified-site 6 /note= "N-terminal acetyl"
FT /note= "(S,S)allothreonyl carbonyl residue forming a
FT keto-amide linkage with residue 7"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 91.7%; Score 44; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXSYS 11
Db |||||
1 EEVVPXGMSYS 11
RESULT 21
ABB80568
ID ABB80568 standard; peptide; 11 AA.
XX
AC ABB80568;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 6 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Alpha-propionyl-glycyl-carbonyl residue forming
FT a keto-amide linkage with residue 7"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN 31-JAN-2002.
PD

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;

Query Match 85.4%; Score 41; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.063; Mismatches 0; Gaps 0;

Matches 9; Conservative 0; Indels 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGTDYS 11

RESULT 24

ABB80551

ID ABB80551 standard; peptide; 11 AA.

XX AC ABB80551;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.4%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.063;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGSDYS 11

RESULT 25

ABB80556

ID ABB80556 standard; peptide; 11 AA.

XX AC ABB80556;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.4%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.063;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

DB 1 EEVVPXGSDYS 11

Search completed: June 10, 2003, 13:39:10
Job time : 32.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds
(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EEVFXGXSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2.6/ptodata/1/iaa/backfilesi.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	66.7	45	2	US-08-637-759B-236
2	32	66.7	45	3	US-08-871-355A-236
3	32	66.7	45	4	US-09-201-945-236
4	32	66.7	947	4	US-09-228-986-73
5	31	64.6	159	2	US-08-844-086-4
6	31	64.6	159	3	US-09-018-211-4
7	31	64.6	513	5	PCT-US91-02714-26
8	31	64.6	622	2	US-08-459-146-2
9	31	64.6	622	1	US-08-459-065-2
10	31	64.6	688	1	US-07-688-352C-28
11	31	64.6	688	2	US-08-474-379C-28
12	31	64.6	688	3	US-09-146-249A-28
13	31	64.6	688	3	US-08-206-188B-28
14	31	64.6	833	2	US-08-844-086-2
15	31	64.6	833	3	US-09-018-211-2
16	30	62.5	121	4	US-09-152-060-85
17	30	62.5	121	4	US-09-152-060-85
18	30	62.5	122	2	US-08-879-995A-1
19	30	62.5	122	3	US-09-215-096-1
20	30	62.5	241	3	US-08-834-776A-2
21	30	62.5	480	1	US-07-803-636A-2
22	30	62.5	626	4	US-08-961-083-220
23	30	62.5	3072	4	US-09-413-814-93
24	30	62.5	3079	4	US-09-413-814-80
25	29	60.4	10	2	US-08-361-517-19
26	29	60.4	10	5	PCT-US93-07964-19
27	29	60.4	20	1	US-08-440-861-53

28	29	60.4	28	1	US-07-971-096-8	Sequence 8, Appl
29	29	60.4	28	1	US-08-175-096-8	Sequence 8, Appl
30	29	60.4	100	2	US-09-047-125-27	Sequence 27, Appl
31	29	60.4	100	3	US-07-736-335B-27	Sequence 27, Appl
32	29	60.4	145	4	US-08-413-974-4	Sequence 4, Appl
33	29	60.4	145	4	US-08-434-418-4	Sequence 4, Appl
34	29	60.4	145	4	US-08-433-288-4	Sequence 4, Appl
35	29	60.4	145	4	US-08-174-739A-4	Sequence 4, Appl
36	29	60.4	145	4	US-08-434-256-4	Sequence 4, Appl
37	29	60.4	263	1	US-07-971-096-4	Sequence 4, Appl
38	29	60.4	263	1	US-08-175-096-4	Sequence 4, Appl
39	29	60.4	263	4	US-08-413-974-6	Sequence 6, Appl
40	29	60.4	263	4	US-08-434-418-6	Sequence 6, Appl
41	29	60.4	263	4	US-08-433-288-6	Sequence 6, Appl
42	29	60.4	263	4	US-08-174-739A-6	Sequence 6, Appl
43	29	60.4	263	4	US-08-434-256-6	Sequence 6, Appl
44	29	60.4	382	3	US-08-582-740-70	Sequence 70, Appl
45	29	60.4	382	4	US-09-109-879-70	Sequence 70, Appl
46	29	60.4	406	3	US-08-582-740-68	Sequence 68, Appl
47	29	60.4	406	4	US-09-109-879-68	Sequence 68, Appl
48	29	60.4	529	4	US-09-240-639-4	Sequence 4, Appl
49	29	60.4	894	2	US-08-867-941-15	Sequence 15, Appl
50	29	60.4	1043	2	US-08-724-354D-4	Sequence 4, Appl
51	29	60.4	1043	3	US-09-270-984A-4	Sequence 4, Appl
52	29	60.4	1118	2	US-08-724-354D-2	Sequence 2, Appl
53	29	60.4	1118	3	US-09-270-984A-2	Sequence 2, Appl
54	29	60.4	2432	4	US-09-074-658-15	Sequence 15, Appl
55	28	58.3	59	4	US-08-963-851-14	Sequence 14, Appl
56	28	58.3	232	1	US-08-468-036-7	Sequence 7, Appl
57	28	58.3	232	2	US-08-376-843-7	Sequence 7, Appl
58	28	58.3	290	4	US-09-370-807-14	Sequence 14, Appl
59	28	58.3	290	4	US-09-921-259-14	Sequence 14, Appl
60	28	58.3	316	4	US-09-325-932A-62	Sequence 62, Appl
61	28	58.3	438	4	US-09-044-718-2	Sequence 2, Appl
62	28	58.3	430	4	US-09-044-718-15	Sequence 15, Appl
63	28	58.3	466	4	US-08-868-435-35	Sequence 35, Appl
64	28	58.3	466	4	US-08-744-231-35	Sequence 35, Appl
65	28	58.3	499	2	US-09-032-315-2	Sequence 2, Appl
66	28	58.3	499	2	US-08-993-318A-2	Sequence 2, Appl
67	28	58.3	499	4	US-09-399-886-2	Sequence 2, Appl
68	28	58.3	499	4	US-09-396-260-2	Sequence 2, Appl
69	28	58.3	499	4	US-09-576-281-2	Sequence 2, Appl
70	28	58.3	520	1	US-08-462-484-2	Sequence 2, Appl
71	28	58.3	520	1	US-08-441-147-2	Sequence 2, Appl
72	28	58.3	520	5	PCT-US95-07536-2	Sequence 2, Appl
73	28	58.3	565	4	US-09-370-807-8	Sequence 8, Appl
74	28	58.3	565	4	US-09-521-259-8	Sequence 8, Appl
75	28	58.3	873	3	US-08-990-140-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-637-759B-236
; Sequence 236, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-236

Query Match 66.7%; Score 32; DB 2; Length 45;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
DB 1 EEISPLGWSY 10

RESULT 2
US-08-871-355A-236
Sequence 236, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-236

Query Match 66.7%; Score 32; DB 3; Length 45;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
DB 1 EEISPLGWSY 10

RESULT 3
US-09-201-945-236
Sequence 236, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-236

Query Match 66.7%; Score 32; DB 4; Length 45;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
DB 1 EEISPLGWSY 10

RESULT 4

US-09-228-986-73
; Sequence 73, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-73

Query Match 66.7%; Score 32; DB 4; Length 947;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 11
|:| | | |
Db 686 VMPGSGISY 694

RESULT 5
US-08-844-086-4
; Sequence 4, Application US/08844086
; Patent No. 5866390
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5866390el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,086
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-4
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-844-086-4

Query Match 64.6%; Score 31; DB 2; Length 159;

Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 BEVVPXGXS 9
|:| | | |
Db 123 BEVLDPGTS 131

RESULT 6
US-09-018-211-4
; Sequence 4, Application US/09018211
; Patent No. 6048716
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6048716el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,211
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,086
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-018-211-4

Query Match 64.6%; Score 31; DB 3; Length 159;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGXS 9
|:| | | |
Db 123 BEVLDPGTS 131

RESULT 7
PCT-US91-02714-26
; Sequence 26, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02714
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-02714-26

Query Match 64.6%; Score 31; DB 5; Length 513;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 10
||| | | |
Db 201 VVPAGSY 208

RESULT 8
US-08-459-146-2
; Sequence 2, Application US/08459146
; Patent No. 5866405
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,146
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117

; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryphonectria
; ORGANISM: parasitica)
; STRAIN: EP713
; US-08-459-146-2

Query Match 64.6%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | | |
Db 31 EEVVPAG 37

RESULT 9
US-08-459-065-2
; Sequence 2, Application US/08459065
; Patent No. 5882642
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; APPLICANT: Nuss, Donald Lee
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,065
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

;; ORIGINAL SOURCE:
;; ORGANISM: Endothia parasitica (Cryphonectria
;; ORGANISM: parasitica)
;; STRAIN: EP713
US-08-459-065-2

Query Match 64.6%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 31 EEVVPAG 37

RESULT 10
US-07-688-352C-28
; Sequence 28, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/688,352C
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-688-352C-28

Query Match 64.6%; Score 31; DB 1; Length 688;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
Db 201 VVPAGGSY 208

RESULT 11
US-08-474-379C-28

; Sequence 28, Application US/08474379C
; Patent No. 5977305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188
FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-379C-28

Query Match 64.6%; Score 31; DB 2; Length 688;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
Db 201 VVPAGGSY 208

RESULT 12
US-09-146-249A-28
; Sequence 28, Application US/09146249A
; Patent No. 6069240
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America

;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/146,249A
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/511,715
;; FILING DATE: 20-APR-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clough, David W.
;; REGISTRATION NUMBER: 36,107
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/474-6300
;; TELEFAX: 312/474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 688 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-146-249A-28

Query Match 64.6%; Score 31; DB 3; Length 688;
Best Local Similarity 75.0%; Pred. NO. 2.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | |
Db 201 VVPAGGSY 208

RESULT 13
US-08-206-188B-28
; Sequence 28, Application US/08206188B
; Patent No. 6100025
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,188B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 688 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-206-188B-28

Query Match 64.6%; Score 31; DB 3; Length 688;
Best Local Similarity 75.0%; Pred. NO. 2.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | |
Db 201 VVPAGGSY 208

RESULT 14
US-08-844-086-2
; Sequence 2, Application US/08844086
; Patent No. 5866390
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NO. 5866390el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,086
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 833 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-844-086-2

Query Match 64.6%; Score 31; DB 2; Length 833;
Best Local Similarity 66.7%; Pred. NO. 2.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXS 9
||| | |
Db 169 EEVLPGTS 177

RESULT 15
US-09-018-211-2
; Sequence 2, Application US/09018211

Patent No. 6048716
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6048716el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,211
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,086
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: F31457-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-018-211-2

Query Match 64.6%; Score 31; DB 3; Length 833;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGS 9
| | | | |
Db 169 EEVLPGDGS 177

RESULT 16
US-09-152-060-68
; Sequence 68, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-152-060-85

EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-152-060-68

Query Match 62.5%; Score 30; DB 4; Length 121;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

RESULT 17
US-09-152-060-85
; Sequence 85, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-152-060-85

Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVPPG 34

RESULT 18
US-08-879-995A-1
; Sequence 1, Application US/08879995A
; Patent No. 5985606
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: Herewith
; APPLICATION NUMBER: US/08/879,995A
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2109906
; TELEFAX: 415-845-4166
; TELEX:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2109906
; US-08-879-995A-1

Query Match 62.5%; Score 30; DB 2; Length 122;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVPPG 34

RESULT 19
US-09-215-096-1
; Sequence 1, Application US/09215096
; Patent No. 6008194
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.

; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/879,995
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2109906
; US-09-215-096-1

Query Match 62.5%; Score 30; DB 3; Length 122;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVPPG 34

RESULT 20
US-08-834-776A-2
; Sequence 2, Application US/08834776A
; Patent No. 6060241
; GENERAL INFORMATION:
; APPLICANT: Cortesey-Theulaz, Irene
; TITLE OF INVENTION: Compositions and Methods Relating to
; TITLE OF INVENTION: Drug Discovery and Detection and Treatment of
; TITLE OF INVENTION: Gastrointestinal Diseases
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,776A

```

QY      1 EVVPXGXS 9
      |||||
Db      407 EETVPSGDS 415

RESULT 22
US-08-961-083-220
; Sequence 220, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 220:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 626 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-220

Query Match 62.5%; Score 30; DB 4; Length 626;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGXS 10
      |||||
Db      70 VVPNGKSY 77

RESULT 23
US-09-413-814-93
; Sequence 93, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans

```

```
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413.814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 3072
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-93

Query Match      62.5%; Score 30; DB 4; Length 3072;
Best Local Similarity 36.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGXSYS 11
Db      2013 EQLPGGTAYN 2023

RESULT 24
US-09-413-814-80
; Sequence 80, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413.814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 3079
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-80

Query Match      62.5%; Score 30; DB 4; Length 3079;
Best Local Similarity 36.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGXSYS 11
Db      2019 EQLPGGTAYN 2029

RESULT 25
US-08-361-517-19
; Sequence 19, Application US/08361517
; Patent No. 5916876
; GENERAL INFORMATION:
; APPLICANT: George A. Heavner and Leon A. Epps
; TITLE OF INVENTION: Peptide Inhibitors of Leukocyte
; TITLE OF INVENTION: Adhesion
; NUMBER OF SEQUENCES: 42
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz
; ADDRESSEE: and No. 5916876ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,517
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/941,652
; FILING DATE: September 8, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: CCOR-0209
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acid residues
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE:
; DESCRIPTION: Amide terminated
US-08-361-517-19

Query Match      60.4%; Score 29; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 5;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EEVVPXGXSY 10
Db      1 KVLPGSSY 9

Search completed: June 10, 2003, 13:51:35
Job time : 10.6429 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds
(without alignments)
75.710 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EEWVPXGXSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	75.0	3472	9 US-10-027-806-4	Sequence 4, Appli
2	36	75.0	3472	9 US-10-034-623-4	Sequence 4, Appli
3	36	75.0	3472	9 US-10-027-801-4	Sequence 4, Appli
4	34	70.8	478	9 US-09-924-340-108	Sequence 108, App
5	34	70.8	478	9 US-09-992-600A-108	Sequence 108, App
6	34	70.8	478	9 US-09-746-783-184	Sequence 184, App
7	34	70.8	478	9 US-10-000-489-108	Sequence 108, App
8	34	70.8	478	9 US-10-000-986-108	Sequence 108, App
9	32	66.7	426	9 US-10-214-766-43	Sequence 43, Appli
10	32	66.7	947	9 US-10-101-464A-73	Sequence 73, Appli
11	31	64.6	597	10 US-09-815-242-13226	Sequence 13226, A
12	31	64.6	653	9 US-09-820-843A-26	Sequence 26, Appli
13	31	64.6	1083	9 US-09-835-976B-2	Sequence 2, Appli
14	31	64.6	1083	9 US-09-835-976B-14	Sequence 14, Appli
15	31	64.6	1099	9 US-09-835-976B-10	Sequence 10, Appli
16	31	64.6	1116	9 US-09-835-976B-12	Sequence 12, Appli
17	31	64.6	1135	9 US-09-835-976B-4	Sequence 4, Appli
18	31	64.6	1135	9 US-09-835-976B-6	Sequence 6, Appli
19	31	64.6	1150	9 US-09-835-976B-8	Sequence 8, Appli

RESULT 1

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOOP-002A
; CURRENT APPLICATION NUMBER: US/10/027.806
; CURRENT FILING DATE: 2001-12-21

ALIGNMENTS

31	64.6	1150	9	US-09-835-976B-16	Sequence 16, Appli
20	62.5	7	9	US-09-909-062-1	Sequence 1, Appli
21	62.5	7	9	US-09-909-062-9	Sequence 9, Appli
22	62.5	7	9	US-09-909-062-130	Sequence 130, App
23	62.5	121	9	US-09-852-797-68	Sequence 68, Appli
24	62.5	121	9	US-09-852-797-85	Sequence 85, Appli
25	62.5	121	10	US-09-853-161-68	Sequence 68, Appli
26	62.5	121	10	US-09-853-161-85	Sequence 85, Appli
27	62.5	121	10	US-09-852-659A-68	Sequence 68, Appli
28	62.5	121	10	US-09-852-659A-85	Sequence 85, Appli
29	62.5	135	9	US-09-992-598-359	Sequence 359, App
30	62.5	135	9	US-09-989-293A-359	Sequence 359, App
31	62.5	135	9	US-09-989-735-359	Sequence 359, App
32	62.5	135	9	US-09-989-735-359	Sequence 359, App
33	62.5	135	9	US-09-989-730-359	Sequence 359, App
34	62.5	135	9	US-09-989-730-359	Sequence 359, App
35	62.5	135	9	US-09-990-436-359	Sequence 359, App
36	62.5	135	9	US-09-991-181-359	Sequence 359, App
37	62.5	135	9	US-09-993-687-359	Sequence 359, App
38	62.5	135	9	US-09-989-734-359	Sequence 359, App
39	62.5	135	9	US-09-997-653-359	Sequence 359, App
40	62.5	135	9	US-10-174-590-444	Sequence 444, App
41	62.5	135	9	US-10-176-758-444	Sequence 444, App
42	62.5	135	9	US-10-175-737-444	Sequence 444, App
43	62.5	135	9	US-09-993-667-359	Sequence 359, App
44	62.5	135	9	US-10-173-706-444	Sequence 444, App
45	62.5	135	9	US-10-175-738-444	Sequence 444, App
46	62.5	135	9	US-10-175-752-444	Sequence 444, App
47	62.5	135	9	US-10-176-482-444	Sequence 444, App
48	62.5	135	9	US-10-176-757-444	Sequence 444, App
49	62.5	135	9	US-10-176-913-444	Sequence 444, App
50	62.5	135	9	US-10-180-552-444	Sequence 444, App
51	62.5	135	9	US-10-180-557-444	Sequence 444, App
52	62.5	135	9	US-09-990-438-359	Sequence 359, App
53	62.5	135	9	US-09-990-562-359	Sequence 359, App
54	62.5	135	9	US-09-997-428-359	Sequence 359, App
55	62.5	135	9	US-09-997-666-359	Sequence 359, App
56	62.5	135	9	US-10-173-700-444	Sequence 444, App
57	62.5	135	9	US-10-174-572-444	Sequence 444, App
58	62.5	135	9	US-10-174-579-444	Sequence 444, App
59	62.5	135	9	US-10-174-582-444	Sequence 444, App
60	62.5	135	9	US-10-174-588-444	Sequence 444, App
61	62.5	135	9	US-10-175-739-444	Sequence 444, App
62	62.5	135	9	US-10-175-740-444	Sequence 444, App
63	62.5	135	9	US-10-175-743-444	Sequence 444, App
64	62.5	135	9	US-10-176-488-444	Sequence 444, App
65	62.5	135	9	US-10-176-492-444	Sequence 444, App
66	62.5	135	9	US-10-176-747-444	Sequence 444, App
67	62.5	135	9	US-10-176-750-444	Sequence 444, App
68	62.5	135	9	US-10-176-985-444	Sequence 444, App
69	62.5	135	9	US-10-176-987-444	Sequence 444, App
70	62.5	135	9	US-10-176-991-444	Sequence 444, App
71	62.5	135	9	US-10-176-992-444	Sequence 444, App
72	62.5	135	9	US-10-176-993-444	Sequence 444, App
73	62.5	135	9	US-10-184-658-444	Sequence 108, App
74	62.5	135	9	US-10-227-884-108	Sequence 359, App
75	62.5	135	9	US-09-990-711-359	

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 75.0%; Score 36; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
|:|:| | |:
DB 2294 EDVIPRGISFS 2304

RESULT 2
US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/034,623
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 75.0%; Score 36; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
|:|:| | |:
DB 2294 EDVIPRGISFS 2304

RESULT 3
US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/027,801
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match 75.0%; Score 36; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
|:|:| | |:
DB 2294 EDVIPRGISFS 2304

RESULT 4
US-09-924-340-108
; Sequence 108, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-108

Query Match 70.8%; Score 34; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EEVVPXGXSYS 11
|:|:| | |:
DB 239 EVAPAGASYN 248

RESULT 5
US-09-992-600A-108
; Sequence 108, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-108

```

Query Match 70.8%; Score 34; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 3; Indels

Qy 2 EVVPXGXSYS 11
 || | | | | :
Db 239 EVAPAGASYN 248

RESULT 6

US-09-746-783-184
; Sequence 184, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth

Query Match 70.8%; Score 34; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 3; Indels

QY 2 EVVPXGXSYS 11
 111111:
 Db 239 EVAPAGASYN 248

RESULT 7
US-10-000-489-108

; Sequence 108, Application US/10000489
; Publication No. US20030092011A1
GENERAL INFORMATION:

Query Match 70.8%; Score 34; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6: Conservative 1; Mismatches 3; Indels

```
Qy      2 EVVPXGXSYS 11
         |||||
Db     239 EVAPAGASYN 248
```

RESULT 8

US-10-000-986-108
; Sequence 108, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

Query Match 70.8%; Score 34; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6: Conservative 1; Mismatches 3; Indels

QY 2 EVVPXGXSYS 11
|||
Db 239 EVAPAGASYN 248

RESULT 9

US-10-214-766-43
; Sequence 43, Application US/10214766
; Publication No. US20030084473A1
; GENERAL INFORMATION:
; APPLICANT: Gocal, Greg
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
; FILE REFERENCE: CA1138
; CURRENT APPLICATION NUMBER: US/10/214,766
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,734
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-214-766-43

Query Match 66.7%; Score 32; DB 9; Length 426;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGXSYS 10
|||
Db 223 EVIPAGQSY 232

RESULT 10

US-10-101-464A-73
; Sequence 73, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-73

Query Match 66.7%; Score 32; DB 9; Length 947;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSYS 11
|||
Db 686 VMPGSGISYS 694

RESULT 11

Query Match 64.6%; Score 31; DB 9; Length 653;

US-09-815-242-13226
; Sequence 13226, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13226
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13226

Query Match 64.6%; Score 31; DB 10; Length 597;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGXS 9
|||
Db 169 EEVLPDGT 177

RESULT 12

US-09-820-843A-26
; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE P
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Vibrio cholerae
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Iron(III) ABC transporter, permease protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: g1|9654609
US-09-820-843A-26

Query Match

Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 300 EEVVPXG 306

RESULT 13
US-09-835-976B-2
; Sequence 2, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (86)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1028)
; OTHER INFORMATION: Xaa-Leu or Ile
US-09-835-976B-2

Query Match 64.6%; Score 31; DB 9; Length 1083;
Best Local Similarity 75.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 10
| | | | |
Db 179 VVPXGXSY 186

RESULT 14
US-09-835-976B-14
; Sequence 14, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: mouse

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (71)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (155)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (467)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (586)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (639)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1023)
; OTHER INFORMATION: Xaa-Leu or Ile
US-09-835-976B-14

Query Match 64.6%; Score 31; DB 9; Length 1083;
Best Local Similarity 75.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 10
| | | | |
Db 179 VVPXGXSY 186

RESULT 15
US-09-835-976B-10
; Sequence 10, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1099
; TYPE: PRT
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (537)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (974)
; OTHER INFORMATION: n-c or a, Xaa-Leu or Ile
US-09-835-976B-10

Query Match 64.6%; Score 31; DB 9; Length 1099;
Best Local Similarity 75.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | ||
Db 195 VVPAGGSY 202

RESULT 16

US-09-835-976B-12

; Sequence 12, Application US/09835976B

; Publication No. US20030027983A1

; GENERAL INFORMATION:

; APPLICANT: Mount, David B.

; APPLICANT: Delpire, Eric

; APPLICANT: Gamba, Gerardo

; APPLICANT: Alfred L. George, Jr.

; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC

; TITLE OF INVENTION: POLYPEPTIDES AND

; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME

; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2

; CURRENT APPLICATION NUMBER: US/09/835,976B

; CURRENT FILING DATE: 2001-04-16

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12

; LENGTH: 1116

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (120)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (358)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (916)

; OTHER INFORMATION: Xaa=Leu or Ile

US-09-835-976B-12

Query Match

Best Local Similarity 64.6%; Score 31; DB 9; Length 1116;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | ||

Db 158 VVPAGGSY 165

RESULT 17

US-09-835-976B-4

; Sequence 4, Application US/09835976B

; Publication No. US20030027983A1

; GENERAL INFORMATION:

; APPLICANT: Mount, David B.

; APPLICANT: Delpire, Eric

; APPLICANT: Gamba, Gerardo

; APPLICANT: Alfred L. George, Jr.

; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC

; TITLE OF INVENTION: POLYPEPTIDES AND

; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME

; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2

; CURRENT APPLICATION NUMBER: US/09/835,976B

; CURRENT FILING DATE: 2001-04-16

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1135

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (35)
; OTHER INFORMATION: Xaa=Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (472)
; OTHER INFORMATION: Xaa=Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1075)
; OTHER INFORMATION: Xaa=Leu or Ile
US-09-835-976B-4

Query Match 64.6%; Score 31; DB 9; Length 1135;

Best Local Similarity 75.0%; Pred. No. 5.1e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | ||

Db 231 VVPAGGSY 238

RESULT 18

US-09-835-976B-6

; Sequence 6, Application US/09835976B

; Publication No. US20030027983A1

; GENERAL INFORMATION:

; APPLICANT: Mount, David B.

; APPLICANT: Delpire, Eric

; APPLICANT: Gamba, Gerardo

; APPLICANT: Alfred L. George, Jr.

; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC

; TITLE OF INVENTION: POLYPEPTIDES AND

; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME

; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2

; CURRENT APPLICATION NUMBER: US/09/835,976B

; CURRENT FILING DATE: 2001-04-16

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 1135

; TYPE: PRT

; ORGANISM: mouse

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (102)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (631)

; OTHER INFORMATION: Xaa=Leu or Ile

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1113)

; OTHER INFORMATION: Xaa=Leu or Ile

US-09-835-976B-6

Query Match

Best Local Similarity 64.6%; Score 31; DB 9; Length 1135;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
||| | ||

Db 231 VVPAGGSY 238

RESULT 19

US-09-835-976B-8

; Sequence 8, Application US/09835976B

; Publication No. US20030027983A1

; GENERAL INFORMATION:

; APPLICANT: Mount, David B.

; APPLICANT: Delpire, Eric

; APPLICANT: Gamba, Gerardo

; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (35)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (308)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (848)
; OTHER INFORMATION: Xaa-Leu or Ile
US-09-835-976B-8

Query Match 64.6%; Score 31; DB 9; Length 1150;
Best Local Similarity 75.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 10
||| |||
Db 246 VVPAGGSY 253

RESULT 20
US-09-835-976B-16
; Sequence 16, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (35)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (346)
; OTHER INFORMATION: Xaa-Leu or Ile
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (789)
; OTHER INFORMATION: Xaa-Leu or Ile
US-09-835-976B-16

Query Match 64.6%; Score 31; DB 9; Length 1150;
Best Local Similarity 75.0%; Pred. No. 5.2e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 VVPXGXSY 10
||| |||
Db 246 VVPAGGSY 253

RESULT 21

US-09-909-062-1
; Sequence 1, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavaliabhan, Viyyor M
; APPLICANT: Lovey, Raymond G
; APPLICANT: Jao, Edwin
; APPLICANT: Bennett, Frank
; APPLICANT: McCormick, Jinping L
; APPLICANT: Pike, Russell E
; APPLICANT: Bogen, Stephane L
; APPLICANT: Liu, Yi-Tsung
; APPLICANT: Arasappan, Ashok
; APPLICANT: Pinto, Patrick A
; APPLICANT: Njoroge, F George
; APPLICANT: Ganguly, Ashit
; TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS
; FILE REFERENCE: IN01157K-US
; CURRENT APPLICATION NUMBER: US/09/909,062
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,109
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLATION
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: norvaline-C(=O)
US-09-909-062-1

Query Match 62.5%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
|||||
Db 1 EEVVPXG 7

RESULT 22

US-09-909-062-9
; Sequence 9, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavaliabhan, Viyyor M
; APPLICANT: Lovey, Raymond G
; APPLICANT: Jao, Edwin
; APPLICANT: Bennett, Frank
; APPLICANT: McCormick, Jinping L
; APPLICANT: Pike, Russell E
; APPLICANT: Bogen, Stephane L
; APPLICANT: Liu, Yi-Tsung
; APPLICANT: Arasappan, Ashok

```

; NAME/KEY: MOD_RES
; LOCATION: (7)..(7)
; OTHER INFORMATION: AMIDATION
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: norvaline-(C=O)
; US-09-909-062-130

Query Match
Best Local Similarity 62.5%; Score 30; DB 9; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 1 EEVVPXG 7

RESULT 24
US-09-852-797-68
; Sequence 68, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: F2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-852-797-68

Query Match
Best Local Similarity 62.5%; Score 30; DB 9; Length 121;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVVPXG 34

RESULT 25
US-09-852-797-85
; Sequence 85, Application US/09852797
; Patent No. US20020172994A1

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003p2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-797-85

Query Match 62.5%; Score 30; DB 9; Length 121;
Best Local Similarity 85.7%; Pred. NO. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVPPGG 34

Search completed: June 10, 2003, 14:35:45
Job time : 16.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11:2143 Seconds
(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EEVVPXGXSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	75.0	3472	2 T31308	hypothetical 367K
2	35	72.9	225	2 S57810	hypothetical prote
3	34	70.8	102	2 A42452	VI protein - tobac
4	34	70.8	165	2 D69493	hypothetical prote
5	34	70.8	259	2 T34536	hypothetical prote
6	34	70.8	1028	2 AF3286	ATP-dependent DNA
7	33	68.8	124	1 VKLJSL	trans-regulatory s
8	33	68.8	427	2 F64084	tolB protein - Hae
9	32	66.7	227	2 E75619	hypothetical prote
10	32	66.7	425	2 T24111	hypothetical prote
11	32	66.7	426	2 D82163	3-phosphoshikimate
12	32	66.7	670	2 S22293	zinc finger protei
13	32	66.7	890	2 A30481	bacteriocin BCN5 -
14	32	66.7	2717	2 A34203	DNA-binding protei
15	31	64.6	123	2 E69342	conserved hypotet
16	31	64.6	284	2 S75817	hypothetical prote
17	31	64.6	319	2 S03833	hypothetical prote
18	31	64.6	361	2 S15299	dtppglucose 4,6-de
19	31	64.6	361	2 AF0767	hypothetical prote
20	31	64.6	437	2 D2945	hypothetical prote
21	31	64.6	450	2 C98337	periplasmic sorbit
22	31	64.6	541	2 AH2679	ATP-dependent DNA
23	31	64.6	544	2 C82900	probable ABC subst
24	31	64.6	561	2 C84239	type II secretion
25	31	64.6	573	2 P97461	DNA ligase (AB0425
26	31	64.6	612	2 T05331	hypothetical prote
27	31	64.6	622	2 S15009	hypothetical prote
28	31	64.6	646	2 C95978	probable beta-gala
29	31	64.6	653	2 D82352	iron(III) ABC tran

30	31	64.6	701	2 S61239	hypothetical prote
31	31	64.6	829	2 H86726	leucine-tRNA ligas
32	31	64.6	833	2 H95029	leucyl-tRNA synth
33	31	64.6	833	2 T39116	leucine-tRNA ligas
34	31	64.6	840	2 T39116	probable sulfate p
35	31	64.6	877	2 T40413	sulfate permease -
36	31	64.6	926	2 AG1860	hypothetical prote
37	31	64.6	1014	2 T17275	hypothetical prote
38	31	64.6	1081	1 A42399	isoleucine-tRNA li
39	31	64.6	1085	2 T18369	K-Cl cotransport p
40	31	64.6	1085	2 T31429	K-Cl cotransport p
41	31	64.6	1086	2 T14114	K-Cl cotransport p
42	31	64.6	1088	2 D82246	probable chitinase
43	31	64.6	1116	2 T31432	K-Cl cotransport p
44	31	64.6	1152	2 D87046	conserved hypotet
45	31	64.6	1548	2 T04456	hypothetical prote
46	30	62.5	84	2 E97333	hypothetical prote
47	30	62.5	118	2 F84263	transcription repr
48	30	62.5	163	2 F97332	hypothetical prote
49	30	62.5	175	2 P0616	transport protein
50	30	62.5	192	2 D83632	hypothetical prote
51	30	62.5	224	2 G70709	probable purQ prot
52	30	62.5	224	2 F87186	phosphoribosylform
53	30	62.5	224	2 D84345	phosphoribosylform
54	30	62.5	227	2 C71211	probable phosphori
55	30	62.5	232	2 C71908	3-oxoacid CoA-tran
56	30	62.5	232	2 C64606	3-oxoacid CoA-t
57	30	62.5	234	2 I38667	melanoma antigen M
58	30	62.5	298	2 T47670	beta-ketoacyl-ACP
59	30	62.5	299	2 C97102	hypothetical prote
60	30	62.5	315	2 I38668	melanoma antigen M
61	30	62.5	327	2 D71278	hypothetical prote
62	30	62.5	363	2 G95237	conserved hypotet
63	30	62.5	363	2 H98101	conserved hypotet
64	30	62.5	363	2 D69551	conserved hypotet
65	30	62.5	404	2 S57178	2-nitropropene dio
66	30	62.5	430	2 B40646	folyl-polyglutam
67	30	62.5	435	2 T46443	hypothetical prote
68	30	62.5	449	2 A99286	conserved hypotet
69	30	62.5	454	2 F83760	hypothetical prote
70	30	62.5	480	1 A45614	merozoite surface
71	30	62.5	480	2 D45661	merozoite surface
72	30	62.5	498	2 B90604	conserved hypotet
73	30	62.5	526	2 T13687	hypothetical prote
74	30	62.5	540	2 E95260	ABC transporter, A
75	30	62.5	540	2 H98125	hypothetical prote

ALIGNMENTS

RESULT 1

T31308
hypothetical 367K protein - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T31308
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of
A:Reference number: Z20994; MUID:98422450; PMID:9748430
A:Accession: T31308
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3472 <SCH>
A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AAC62699.1
C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 75.0%; Score 36; DB 2; Length 3472;
Best Local Similarity 54.5%; Pred. No. 79;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11

Db 2294 EDVIPRGLSFS 2304

RESULT 2

S57810

hypothetical protein precursor (clone TPP11) - tomato

C:Species: Lycopersicon esculentum (tomato)

C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000

C:Accession: S57810

R:Milligan, S.B.; Gasser, C.S.

Plant Mol. Biol. 28, 691-711, 1995

A:Title: Nature and regulation of pistil-expressed genes in tomato.

A:Reference number: S57808; MUID:93375233; PMID:7647301

A:Accession: S57810

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-225 <MIL>

A:Cross-references: EMBL:U20592; NID:g924625; PIDN:AAA80497.1; PID:g924626

C:Superfamily: plant Kunitz-type proteinase inhibitor

Query Match

Best Local Similarity 72.9%; Score 35; DB 2; Length 225;

Mismatches 3; Conservative 3; Indels 0; Gaps 0;

Matches 6; Gaps 0;

QY 1 EEVVPXGXSYS 11

Db 32 DEVVPGNKTYA 42

RESULT 3

A42452

V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match

Best Local Similarity 70.8%; Score 34; DB 2; Length 102;

Mismatches 2; Conservative 6; Indels 0; Gaps 0;

Matches 6; Gaps 0;

QY 2 EEVVPXGXSYS 11

Db 7 QVVPSPGINYS 16

RESULT 4

D69493

hypothetical protein AFI949 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 04-Mar-2000

C:Accession: D69493

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69493

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-165 <KLE>

A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AAB89307.1; PID:g264859

C:Superfamily: Archaeoglobus fulgidus hypothetical protein AFI949

Query Match 70.8%; Score 34; DB 2; Length 165;

Best Local Similarity 60.0%; Pred. No. 8.2;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSY 10

Db 60 EESIPDGASY 69

RESULT 5

T34536

hypothetical protein DKFp434C031.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34536

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34536

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-259 <POU>

A:Cross-references: EMBL:AL122063

A:Experimental source: adult testis; clone DKFp434C031

C:Genetics:

A:Note: DKFp434C031.1

Query Match

Best Local Similarity 70.8%; Score 34; DB 2; Length 259;

Mismatches 1; Conservative 6; Indels 0; Gaps 0;

Matches 6; Gaps 0;

QY 2 EWVPXGXSYS 11

Db 22 EVAPAGASYN 31

RESULT 6

AF3286

ATP-dependent DNA helicase BMEI0275 [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AF3286

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivan

; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; L

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella mel

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3286

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1028 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51457.1; PID:g17982167; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0275

A:Map position: 1

Query Match

Best Local Similarity 70.8%; Score 34; DB 2; Length 1028;

Mismatches 2; Conservative 6; Indels 0; Gaps 0;

Matches 6; Gaps 0;

QY 1 EEVVPXGXSYS 11

Db 76 EKIVPPGARYS 86

RESULT 7

VKLJSI

trans-regulatory splicing protein - simian immunodeficiency virus SIVcpz

N:Alternate names: anti-repression trans-activator; art protein; rev protein; trs

C:Species: simian immunodeficiency virus SIVcpz

A:Note: host Pan troglodytes (chimpanzee)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
 C:Accession: S09988
 R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
 Nature 345, 356-359, 1990
 A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
 A:Reference number: S09983; MUID:90259077; PMID:2188136
 A:Accession: S09988
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-124 <HUE>
 A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36405.1; PID:g763085
 C:Genetics:
 A:Gene: rev; trs; art
 A:Introns: 27/1
 C:Superfamily: AIDS trans-regulatory splicing protein
 C:Keywords: AIDS; immunodeficiency; splicing protein; transcription regulation

Query Match 68.8%; Score 33; DB 1; Length 124;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 ||| |
 Db 107 ETVPAGNGYS 116

RESULT 8
 F64064
 tolB protein - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C:Accession: F64064; JCS213
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.
 Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: F64064
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-427 <TTGR>
 A:Cross-references: GB:U32722; GB:I42023; NID:g1573348; PIDN:AAC22040.1; PID:g1573352;
 A:Experimental source: strain Rd KW20
 R:Sen, K.; Sikkema, D.J.; Murphy, T.F.
 Gene 178, 75-81, 1996
 A:Title: Isolation and characterization of the Haemophilus influenzae tolQ, tolR, tolA
 A:Reference number: JCS212; MUID:97080550; PMID:8921895
 A:Accession: JCS213
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-5, 'H', 'I', 'V', '15-16, 'ITH', '20, 'V', '22-78, 'H', '80-128, 'A', '130-159, 'G', '161-236
 A:Cross-references: GB:U32470; NID:g1685076; PIDN:AAC44597.1; PID:g1685080
 A:Experimental source: strain 1479
 C:Genetics:
 A:Gene: tolB
 C:Function:
 A:Description: involved in transport of colicins and phages across the cell envelope; ph

Query Match 68.8%; Score 33; DB 2; Length 427;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 ||| |
 Db 103 QVVPSCNGYS 112

RESULT 9
 E75619
 hypothetical protein DRB0013 - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: E75619
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.
 M.; Shen, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.M.
 S.; Smith, M.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: E75619
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-227 <WHI>
 A:Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12657.1; PID:g6460953; TIGR:
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRB0013
 A:Map position: megaplasmid
 A:Genome: plasmid
 A:Note: plasmid MP1
 C:Superfamily: Deinococcus radiodurans megaplasmid hypothetical protein DRB0013

Query Match 66.7%; Score 32; DB 2; Length 227;
 Best Local Similarity 54.5%; Pred. No. 31;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
 ||| |
 Db 43 ESVLPIGHSFS 53

RESULT 10
 T24111
 hypothetical protein R10D12.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24111
 R:Percy, C.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19842
 A:Accession: T24111
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-425 <WIL>
 A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN000023; CESP:R10D12.10
 A:Experimental source: clone R10D12
 C:Genetics:
 A:Gene: CESP:R10D12.10
 A:Map position: 5
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 66.7%; Score 32; DB 2; Length 425;
 Best Local Similarity 50.0%; Pred. No. 62;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 10
 ||| |
 Db 335 EQIVPGGLQY 344

RESULT 11
 D82163
 3-phosphoshikimate 1-carboxyvinyltransferase VC1732 [imported] - Vibrio cholerae
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: D82163
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.J.; Bass, S.; Qin, H.; Dragoi, I.; Sell
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: D82163

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <HEI>
A:Cross-references: GB:AE004251; GB:AE003852; NID:99656248; PIDN:AAF94882.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1732
A:Map position: 1
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxyvinyltransferase

Query Match 66.7%; Score 32; DB 2; Length 426;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSY 10
| | | | |
DB 223 EFVIPAGQSY 232

RESULT 12
S22293
zinc finger protein AT-BP2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Sep-1999
C:Accession: S22293; I78656
R:Mitchellmore, C.; Traboni, C.; Cortese, R.
Nucleic Acids Res. 19, 141-147, 1991
A>Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-
A:Reference number: F58280; MUID:91187610; PMID:1901405
A:Accession: S22293
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-670 <MIT>
A:Cross-references: EMBL:X54250; NID:957519; PIDN:CAA38151.1; PID:957520
A>Note: the authors did not translate the codon for residue 1
C:Superfamily: HIV-EP2 enhancer-binding protein
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 66.7%; Score 32; DB 2; Length 670;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSY 11
| | | | |
DB 376 VVPAGLTYS 384

RESULT 13
A30481
bacteriocin BCN5 - Clostridium perfringens plasmid pIP404
C:Species: Clostridium perfringens
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Nov-1999
C:Accession: A30481; S03779
R:Garnier, T.; Cole, S.T.
J. Bacteriol. 168, 1189-1196, 1986
A>Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens and
A:Reference number: JTO354; MUID:87057020; PMID:2877971
A:Accession: A30481
A:Molecule type: DNA
A:Residues: 1-890 <GAR>
A:Cross-references: GB:M32882; GB:J03309; NID:gl50738; PIDN:AAA98249.1; PID:gl50739
C:Genetics:
A:Gene: bcn
A:Genome: plasmid
C:Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5
C:Keywords: bacteriocin

Query Match 66.7%; Score 32; DB 2; Length 890;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EEVVPXGXSY 10
| | | | |

Db 170 EVVPGGFTY 178

RESULT 14
A34203
DNA-binding protein PRDII-BF1 - human
N:Alternate names: major histocompatibility complex enhancer-binding protein 1
C:Species: Homo sapiens (man)
C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 20-Sep-1999
C:Accession: A34203; A34779
R:Fan, C.M.; Maniatis, T.
Genes Dev. 4, 29-42, 1990
A>Title: A DNA-binding protein containing two widely separated zinc finger motifs
A:Reference number: A34779; MUID:90169514; PMID:2108471
A:Accession: A34203
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2717 <FAN>
A:Cross-references: EMBL:X51435; NID:938017; PIDN:CAA35798.1; PID:938018
R:Baldwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.
Mol. Cell. Biol. 10, 1406-1414, 1990
A>Title: A large protein containing zinc finger domains binds to related sequence
A:Reference number: A34779; MUID:90205817; PMID:2108316
A:Accession: A34779
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 801-1072, 'N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 1609-1670, 'I', 1672-1717, 'I', 1719-1770, 'I', 1772-1817, 'I', 1819-1870, 'I', 1872-1923, 'I', 1925-1976, 'I', 1978-2029, 'I', 2031-2082, 'I', 2084-2135, 'I', 2137-2188, 'I', 2190-2241, 'I', 2243-2294, 'I', 2296-2347, 'I', 2349-2400, 'I', 2402-2453, 'I', 2455-2506, 'I', 2508-2559, 'I', 2561-2612, 'I', 2614-2665, 'I', 2667-2717, 'I', 2719-2770, 'I', 2772-2823, 'I', 2825-2876, 'I', 2878-2929, 'I', 2931-2982, 'I', 2984-3035, 'I', 3037-3088, 'I', 3090-3141, 'I', 3143-3194, 'I', 3196-3247, 'I', 3249-3300, 'I', 3302-3353, 'I', 3355-3406, 'I', 3408-3459, 'I', 3461-3512, 'I', 3514-3565, 'I', 3567-3618, 'I', 3620-3671, 'I', 3673-3724, 'I', 3726-3777, 'I', 3779-3830, 'I', 3832-3883, 'I', 3885-3936, 'I', 3938-3989, 'I', 3991-4042, 'I', 4044-4095, 'I', 4097-4148, 'I', 4150-4201, 'I', 4203-4254, 'I', 4256-4307, 'I', 4309-4360, 'I', 4362-4413, 'I', 4415-4466, 'I', 4468-4519, 'I', 4521-4572, 'I', 4574-4625, 'I', 4627-4678, 'I', 4680-4731, 'I', 4733-4784, 'I', 4786-4837, 'I', 4839-4890, 'I', 4892-4943, 'I', 4945-4996, 'I', 4998-5049, 'I', 5051-5102, 'I', 5104-5155, 'I', 5157-5208, 'I', 5210-5261, 'I', 5263-5314, 'I', 5316-5367, 'I', 5369-5420, 'I', 5422-5473, 'I', 5475-5526, 'I', 5528-5579, 'I', 5581-5632, 'I', 5634-5685, 'I', 5687-5738, 'I', 5740-5791, 'I', 5793-5844, 'I', 5846-5897, 'I', 5899-5950, 'I', 5952-6003, 'I', 6005-6056, 'I', 6058-6109, 'I', 6111-6162, 'I', 6164-6215, 'I', 6217-6268, 'I', 6270-6321, 'I', 6323-6374, 'I', 6376-6427, 'I', 6429-6480, 'I', 6482-6533, 'I', 6535-6586, 'I', 6588-6639, 'I', 6641-6692, 'I', 6694-6745, 'I', 6747-6798, 'I', 6800-6851, 'I', 6853-6904, 'I', 6906-6957, 'I', 6959-7010, 'I', 7012-7063, 'I', 7065-7116, 'I', 7118-7169, 'I', 7171-7222, 'I', 7224-7275, 'I', 7277-7328, 'I', 7330-7381, 'I', 7383-7434, 'I', 7436-7487, 'I', 7489-7540, 'I', 7542-7593, 'I', 7595-7646, 'I', 7648-7699, 'I', 7701-7752, 'I', 7754-7805, 'I', 7807-7858, 'I', 7860-7911, 'I', 7913-7964, 'I', 7966-8017, 'I', 8019-8070, 'I', 8072-8123, 'I', 8125-8176, 'I', 8178-8229, 'I', 8231-8282, 'I', 8284-8335, 'I', 8337-8388, 'I', 8390-8441, 'I', 8443-8494, 'I', 8496-8547, 'I', 8549-8600, 'I', 8602-8653, 'I', 8655-8706, 'I', 8708-8759, 'I', 8761-8812, 'I', 8814-8865, 'I', 8867-8918, 'I', 8920-8971, 'I', 8973-9024, 'I', 9026-9077, 'I', 9079-9130, 'I', 9132-9183, 'I', 9185-9236, 'I', 9238-9289, 'I', 9291-9342, 'I', 9344-9395, 'I', 9397-9448, 'I', 9450-9501, 'I', 9503-9554, 'I', 9556-9607, 'I', 9609-9660, 'I', 9662-9713, 'I', 9715-9766, 'I', 9768-9819, 'I', 9821-9872, 'I', 9874-9925, 'I', 9927-9978, 'I', 9980-10031, 'I', 10033-10084, 'I', 10086-10137, 'I', 10139-10190, 'I', 10192-10243, 'I', 10245-10296, 'I', 10298-10349, 'I', 10351-10402, 'I', 10404-10455, 'I', 10457-10508, 'I', 10510-10561, 'I', 10563-10614, 'I', 10616-10667, 'I', 10669-10720, 'I', 10722-10773, 'I', 10775-10826, 'I', 10828-10879, 'I', 10881-10932, 'I', 10934-10985, 'I', 10987-11038, 'I', 11040-11091, 'I', 11093-11144, 'I', 11146-11197, 'I', 11199-11250, 'I', 11252-11303, 'I', 11305-11356, 'I', 11358-11409, 'I', 11411-11462, 'I', 11464-11515, 'I', 11517-11568, 'I', 11570-11621, 'I', 11623-11674, 'I', 11676-11727, 'I', 11729-11780, 'I', 11782-11833, 'I', 11835-11886, 'I', 11888-11939, 'I', 11941-11992, 'I', 11994-12045, 'I', 12047-12098, 'I', 12100-12151, 'I', 12153-12204, 'I', 12206-12257, 'I', 12259-12310, 'I', 12312-12363, 'I', 12365-12416, 'I', 12418-12469, 'I', 12471-12522, 'I', 12524-12575, 'I', 12577-12628, 'I', 12630-12681, 'I', 12683-12734, 'I', 12736-12787, 'I', 12789-12840, 'I', 12842-12893, 'I', 12895-12946, 'I', 12948-12999, 'I', 13001-13052, 'I', 13054-13105, 'I', 13107-13158, 'I', 13160-13211, 'I', 13213-13264, 'I', 13266-13317, 'I', 13319-13370, 'I', 13372-13423, 'I', 13425-13476, 'I', 13478-13529, 'I', 13531-13582, 'I', 13584-13635, 'I', 13637-13688, 'I', 13690-13741, 'I', 13743-13794, 'I', 13796-13847, 'I', 13849-13900, 'I', 13902-13953, 'I', 13955-14006, 'I', 14008-14059, 'I', 14061-14112, 'I', 14114-14165, 'I', 14167-14218, 'I', 14220-14271, 'I', 14273-14324, 'I', 14326-14377, 'I', 14379-14430, 'I', 14432-14483, 'I', 14485-14536, 'I', 14538-14589, 'I', 14591-14642, 'I', 14644-14695, 'I', 14697-14748, 'I', 14750-14801, 'I', 14803-14854, 'I', 14856-14907, 'I', 14909-14960, 'I', 14962-15013, 'I', 15015-15066, 'I', 15068-15119, 'I', 15121-15172, 'I', 15174-15225, 'I', 15227-15278, 'I', 15280-15331, 'I', 15333-15384, 'I', 15386-15437, 'I', 15439-15490, 'I', 15492-15543, 'I', 15545-15596, 'I', 15598-15649, 'I', 15651-15702, 'I', 15704-15755, 'I', 15757-15808, 'I', 15810-15861, 'I', 15863-15914, 'I', 15916-15967, 'I', 15969-16019, 'I', 16021-16072, 'I', 16074-16125, 'I', 16127-16178, 'I', 16180-16231, 'I', 16233-16284, 'I', 16286-16337, 'I', 16339-16390, 'I', 16392-16443, 'I', 16445-16496, 'I', 16498-16549, 'I', 16551-16602, 'I', 16604-16655, 'I', 16657-16708, 'I', 16710-16761, 'I', 16763-16814, 'I', 16816-16867, 'I', 16869-16920, 'I', 16922-16973, 'I', 16975-17026, 'I', 17028-17079, 'I', 17081-17132, 'I', 17134-17185, 'I', 17187-17238, 'I', 17240-17291, 'I', 17293-17344, 'I', 17346-17397, 'I', 17399-17450, 'I', 17452-17503, 'I', 17505-17556, 'I', 17558-17609, 'I', 17611-17662, 'I', 17664-17715, 'I', 17717-17768, 'I', 17770-17821, 'I', 17823-17874, 'I', 17876-17927, 'I', 17929-17980, 'I', 17982-18033, 'I', 18035-18086, 'I', 18088-18139, 'I', 18141-18192, 'I', 18194-18245, 'I', 18247-18298, 'I', 18300-18351, 'I', 18353-18404, 'I', 18406-18457, 'I', 18459-18510, 'I', 18512-18563, 'I', 18565-18616, 'I', 18618-18669, 'I', 18671-18722, 'I', 18724-18775, 'I', 18777-18828, 'I', 18830-18881, 'I', 18883-18934, 'I', 18936-18987, 'I', 18989-19040, 'I', 19042-19093, 'I', 19095-19146, 'I', 19148-19199, 'I', 19201-19252, 'I', 19254-19305, 'I', 19307-19358, 'I', 19360-19411, 'I', 19413-19464, 'I', 19466-19517, 'I', 19519-19570, 'I', 19572-19623, 'I', 19625-19676, 'I', 19678-19729, 'I', 19731-19782, 'I', 19784-19835, 'I', 19837-19888, 'I', 19890-19941, 'I', 19943-19994, 'I', 19996-20047, 'I', 20049-20100, 'I', 20102-20153, 'I', 20155-20206, 'I', 20208-20259, 'I', 20261-20312, 'I', 20314-20365, 'I', 20367-20418, 'I', 20420-20471, 'I', 20473-20524, 'I', 20526-20577, 'I', 20579-20630, 'I', 20632-20683, 'I', 20685-20736, 'I', 20738-20789, 'I', 20791-20842, 'I', 20844-20895, 'I', 20897-20948, 'I', 20950-20999, 'I', 21001-21052, 'I', 21054-21105, 'I', 21107-21158, 'I', 21160-21211, 'I', 21213-21264, 'I', 21266-21317, 'I', 21319-21370, 'I', 21372-21423, 'I', 21425-21476, 'I', 21478-21529, 'I', 21531-21582, 'I', 21584-21635, 'I', 21637-21688, 'I', 21690-21741, 'I', 21743-21794, 'I', 21796-21847, 'I', 21849-21900, 'I', 21902-21953, 'I', 21955-22006, 'I', 22008-22059, 'I', 22061-22112, 'I', 22114-22165, 'I', 22167-22218, 'I', 22220-22271, 'I', 22273-22324, 'I', 22326-22377, 'I', 22379-22430, 'I', 22432-22483, 'I', 22485-22536, 'I', 22538-22589, 'I', 22591-22642, 'I', 22644-22695, 'I', 22697-22748, 'I', 22750-22801, 'I', 22803-22854, 'I', 22856-22907, 'I', 22909-22960, 'I', 22962-23013, 'I', 23015-23066, 'I', 23068-23119, 'I', 23121-23172, 'I', 23174-23225, 'I', 23227-23278, 'I', 23280-23331, 'I', 23333-23384, 'I', 23386-23437, 'I', 23439-23490, 'I', 23492-23543, 'I', 23545-23596, 'I', 23598-23649, 'I', 23651-23702, 'I', 23704-23755, 'I', 23757-23808, 'I', 23810-23861, 'I', 23863-23914, 'I', 23916-23967, 'I', 23969-24020, 'I', 24022-24073, 'I', 24075-24126, 'I', 24128-24179, 'I', 24181-24232, 'I', 24234-24285, 'I', 24287-24338, 'I', 24340-24391, 'I', 24393-24444, 'I', 24446-24497, 'I', 24499-24550, 'I', 24552-24603, 'I', 24605-24656, 'I', 24658-24709, 'I', 24711-24762, 'I', 24764-24815, 'I', 24817-24868, 'I', 24870-24921, 'I', 24923-24974, 'I', 24976-25027, 'I', 25029-25080, 'I', 25082-25133, 'I', 25135-25186, 'I', 25188-25239, 'I', 25241-25292, 'I', 25294-25345, 'I', 25347-25398, 'I', 25400-25451, 'I', 25453-25504, 'I', 25506-25557, 'I', 25559-25610, 'I', 25612-25663, 'I', 25665-25716, 'I', 25718-25769, 'I', 25771-25822, 'I', 25824-25875, 'I', 25877-25928, 'I', 25930-25981, 'I', 25983-26034, 'I', 26036-26087, 'I', 26089-26140, 'I', 26142-26193, 'I', 26195-26246, 'I', 26248-26299, 'I', 26301-26352, 'I', 26354-26405, 'I', 26407-26458, 'I', 26460-26511, 'I', 26513-26564, 'I', 26566-26617, 'I', 26619-26670, 'I', 26672-26723, 'I', 26725-26776, 'I', 26778-26829, 'I', 26831-26882, 'I', 26884-26935, 'I', 26937-26988, 'I', 26990-27041, 'I', 27043-27094, 'I', 27096-27147, 'I', 27149-27200, 'I', 27202-27253, 'I', 27255-27306, 'I', 27308-27359, 'I', 27361-27412, 'I', 27414-27465, 'I', 27467-27518, 'I', 27520-27571, 'I', 27573-27624, 'I', 27626-27677, 'I', 27679-27730, 'I', 27732-27783, 'I', 27785-27836, 'I', 27838-27889, 'I', 27891-27942, 'I', 27944-27995, 'I', 27997-28048, 'I', 28050-28101, 'I', 28103-28154, 'I', 28156-28207, 'I', 28209-28260, 'I', 28262-28313, 'I', 28315-28366, 'I', 28368-28419, 'I', 28421-28472, 'I', 28474-28525, 'I', 28527-28578, 'I', 28580-28631, 'I', 28633-28684, 'I', 28686-28737, 'I', 28739-28790, 'I', 28792-28843, 'I', 28845-28896, 'I', 28898-28949, 'I', 28951-28999, 'I', 29001-29052, 'I', 29054-29105, 'I', 29107-29158, 'I', 29160-29211, 'I', 29213-29264, 'I', 29266-29317, 'I', 29319-29370, 'I', 29372-29423, 'I', 29425-29476, 'I', 29478-29529, 'I', 29531-29582, 'I', 29584-29635, 'I', 29637-29688, 'I', 29690-29741, 'I', 29743-29794, 'I', 29796-29847, 'I', 29849-29900, 'I', 29902-29953, 'I', 29955-30006, 'I', 30008-30059, 'I', 30061-30112, 'I', 30114-30165, 'I', 30167-30218, 'I', 30220-30271, 'I', 30273-30324, 'I', 30326-30377, 'I', 30379-30430, 'I', 30432-30483, 'I', 30485-30536, 'I', 30538-30589, 'I', 30591-30642, 'I', 30644-30695, 'I', 30697-30748, 'I', 30750-30801, 'I', 30803-30854, 'I', 30856-30907, 'I', 30909-30960, 'I', 30962-31013, 'I', 31015-31066, 'I', 31068-31119, 'I', 31121-31172, 'I', 31174-31225, 'I', 31227-31278, 'I', 31280-31331, 'I', 31333-31384, 'I', 31386-31437, 'I', 31439-31490, 'I', 31492-31543, 'I', 31545-31596, 'I', 31598-31649, 'I', 31651-31702, 'I', 31704-31755, 'I', 31757-31808, 'I', 31810-31861, 'I', 31863-31914, 'I', 31916-31967, 'I', 31969-32020, 'I', 32022-32073, 'I', 32075-32126, 'I', 32128-32179, 'I', 32181-32232, 'I', 32234-32285, 'I', 32287-32338, 'I', 32340-32391, 'I', 32393-32444, 'I', 32446-32497, 'I', 32499-32550, 'I', 32552-32603, 'I', 32605-32656, 'I', 32658-32709, 'I', 32711-32762, 'I', 32764-32815, 'I', 32817-32868, 'I', 32870-32921, 'I', 32923-32974, 'I', 32976-33027, 'I', 33029-33080, 'I', 33082-33133, 'I', 33135-33186, 'I', 33188-33239, 'I', 33241-33292, 'I', 33294-33345, 'I', 33347-33398, 'I', 33400-33451, 'I', 33453-33504, 'I', 33506-33557, 'I', 33559-33610, 'I', 33612-33663, 'I', 33665-33716, 'I', 33718-33769, 'I', 33771-33822, 'I', 33824-33875, 'I', 33877-33928, 'I', 33930-33981, 'I', 33983-34034, 'I', 34036-34087, 'I', 34089-34140, 'I', 34142-34193, 'I', 34195-34246, 'I', 34248-34299, 'I', 34301-34352, 'I', 34354-34405, 'I', 34407-34458, 'I', 34460-34511, 'I', 34513-34564, 'I', 34566-34617, 'I', 34619-34670, 'I', 34672-34723, 'I', 34725-34776, 'I', 34778-34829, 'I', 34831-34882, 'I', 34884-34935, 'I', 34937-34988, 'I', 34990-35041, 'I', 35043-35094, 'I', 35096-35147, 'I', 35149-35200, 'I', 35202-35253, 'I', 35255-35306, 'I', 35308-35359, 'I', 35361-35412, 'I', 35414-35465, 'I', 35467-35518, 'I', 35520-35571, 'I', 35573-35624, 'I', 35626-35677, 'I', 35679-35730, 'I', 35732-35783, 'I', 35785-35836, 'I', 35838-35889, 'I', 35891-35942, 'I', 35944-35995, 'I', 35997-36048, 'I', 36050-36101, 'I', 36103-36154, 'I', 36156-36207, 'I', 36209-36260, 'I', 36262-36313, 'I', 36315-36366, 'I', 36368-36419, 'I', 36421-36472, 'I', 36474-36525, 'I', 36527-36578, 'I', 36580-36631, 'I', 36633-36684, 'I', 36686-36737, 'I', 36739-36790, 'I', 36792-3

C.Species: *Synechocystis* sp.
A.Variety: PCC 6803
C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C.Accession: S75817
O.K.: Okumura, S.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O.K.: Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A.Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75817
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-284 <KAN>
A:Cross-references: EMBL:D90913; GB:AB0011339; NID:g16533348; PIDN:BAAL8276.1; PID:g101900
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 64.6%; Score 31; DB 2; Length 284;
Best Local Similarity 55.6%; Pred. No. 66;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSYS 11
|:| | | | |
DB 208 VIPAGVSYT 216

RESULT 17
S03833
hypothetical protein 1 - chestnut blight fungus
C:Species: *Cryphonectria parasitica*, *Endothia parasitica* (chestnut blight fungus)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Sep-1997
C:Accession: S03833
R:Rae, B.P.; Hillman, B.I.; Tagaglia, J.; Nuss, D.L.
EMBO J. 8, 657-663, 1989
A:Title: Characterization of double-stranded RNA genetic elements associated with biological
A:Reference number: S03833; MUID:89251594; PMID:2721496
A:Accession: S03833
A:Molecule type: DNA
A:Residues: 1-319 <RAE>
A:Cross-references: EMBL:X14524; NID:g2624; PID:g2625
A:Note: the authors translated the codon CAG for residue 156 as Gly

Query Match 64.6%; Score 31; DB 2; Length 319;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | | | |
DB 31 EEVVPAG 37

RESULT 18
S15299
dTPDglucose 4,6-dehydratase (EC 4.2.1.46) - *Salmonella typhimurium*
C:Species: *Salmonella typhimurium*
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Oct-1997
C:Accession: S15299
R:Jiang, X.M.; Neal, B.; Santiago, F.; Lee, S.J.; Romana, L.K.; Reeves, P.R.
Mol. Microbiol. 5, 695-713, 1991
A:Title: Structure and sequence of the rfb (O antigen) gene cluster of *Salmonella serovar*
A:Reference number: S15296; MUID:91260454; PMID:1710759
A:Accession: S15299
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <MOL>
C:Genetics:
A:Gene: rfbB
C:Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: carbon-oxygen lyase; hydro-lyase
F:3-341/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 64.6%; Score 31; DB 2; Length 361;
Best Local Similarity 50.0%; Pred. No. 85;

C98337
 periplasmic sorbitol-binding protein, smcE (AF018073) [imported] - Agrobacterium tumefaciens
 C:Species: Agrobacterium tumefaciens
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: C98337
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:11743194
 A:Accession: C98337
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-450 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK90221.1; PID:g15160234; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L3289
 A:Map position: linear chromosome

Query Match 64.6%; Score 31; DB 2; Length 450;
 Best Local Similarity 60.0%; Pred. NO. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 |||||
 Db 302 EVTPNGSSWS 311

RESULT 22
 AH2679
 ATP-dependent DNA ligase Atu0840 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AH2679
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S.; Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AH2679
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-541 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAU41854.1; PID:g17739214; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0840
 A:Map position: circular chromosome

Query Match 64.6%; Score 31; DB 2; Length 541;
 Best Local Similarity 50.0%; Pred. NO. 1.3e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 10
 |||||
 Db 442 EQLVPVGKAY 451

RESULT 23
 C82900
 probable ABC substrate-binding protein, iron U0359 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: C82900
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H. submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mirror image
 A:Reference number: AB2870
 A:Accession: C82900
 A>Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-544 <GLA>
 A:Cross-references: GB:AE002133; GB:AF222894; NID:g6899339; PIDN:AAF30768.1; GSPDB:GN00169
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: ABCsbp-5; U0359
 A:Genetic code: GGC

Query Match 64.6%; Score 31; DB 2; Length 544;
 Best Local Similarity 70.0%; Pred. NO. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 10
 |||||
 Db 135 EEVVPHYLSY 144

RESULT 24

C84239
 type II secretion system protein [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: C84239
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lapan, J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jung, K.H.; Alam, M.; Freitas, T.; Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.N.
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: AB4160; PMID:20504483; PMID:11016950
 A:Accession: C84239
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-561 <STO>
 A:Cross-references: GB:AE004437; NID:g10580393; PIDN:AAG19279.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: gspE2
 C:Superfamily: conserved hypothetical protein MJ0900

Query Match 64.6%; Score 31; DB 2; Length 561;
 Best Local Similarity 45.5%; Pred. NO. 1.4e+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
 |||||
 Db 251 EEVTPRGSAFT 261

RESULT 25

F97461
 DNA ligase (AB042527) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: F97461
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:11743194
 A:Accession: F97461
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-573 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86647.1; PID:g15155825; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_1536
 A:Map position: circular chromosome

Query Match 64.6%; Score 31; DB 2; Length 573;
 Best Local Similarity 50.0%; Pred. NO. 1.4e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 10

Db 474 EQLVPVGKAY 483

Search completed: June 10, 2003, 13:49:16
Job time : 12.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EEVVPXGXSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	75.0	1499	1 A10C_HUMAN	O60312 homo sapien
2	35	72.9	1058	1 CARB_FUSNN	O8r986 fusobacteri
3	34	70.8	102	1 Y11K_TYDVA	P31619 tobacco yel
4	34	70.8	478	1 GSR2_HUMAN	O9pzm5 homo sapien
5	33	68.8	124	1 REV_SIVCZ	P17280 chimpanzee
6	33	68.8	427	1 TOLB_HAEIN	P44677 haemophilus
7	32	66.7	426	1 AROA_VIBCH	O9krb0 vibrio chol
8	32	66.7	890	1 BCNS_CLOPE	P08696 clostridium
9	32	66.7	2717	1 ZEP1_HUMAN	P15822 homo sapien
10	31	64.6	319	1 VHAL_CRYPA	P10941 cryphonectr
11	31	64.6	361	1 REBB_SALTY	P26391 salmonella
12	31	64.6	507	1 THDI_BURCE	P53607 burkholderi
13	31	64.6	829	1 SYL_LACLA	O9chb6 lactococcus
14	31	64.6	877	1 SULH_SCHPO	O74377 schizosacch
15	31	64.6	1081	1 SVI_TETTH	P36422 tetrahymena
16	30	62.5	121	1 TKNK_HUMAN	O9uhf0 homo sapien
17	30	62.5	223	1 PURQ_PYRHO	O59619 pyrococcus
18	30	62.5	224	1 PURQ_HALNI	O9hnu2 halobacteri
19	30	62.5	224	1 PURQ_MYCLE	O05756 mycobacteri
20	30	62.5	224	1 PURQ_MYCTU	P71841 mycobacteri
21	30	62.5	225	1 PURQ_CORAM	O9rthx0 corynebacte
22	30	62.5	232	1 SCOA_HELPJ	O921e3 helicobacte
23	30	62.5	232	1 SCOA_HELPY	P56006 helicobacte
24	30	62.5	234	1 MAG8_HUMAN	P43361 homo sapien
25	30	62.5	315	1 MAG9_HUMAN	P43362 homo sapien
26	30	62.5	404	1 YJ9N_YEAST	P47177 saccharomyc
27	30	62.5	430	1 FOLC_BACSU	O05865 bacillus su
28	30	62.5	471	1 TMLH_NEUCR	O96ub1 neurospora
29	30	62.5	509	1 CHLB_MESVI	O9mur9 mesostigma
30	30	62.5	726	1 PRTP_HSV60	P52384 human herpe
31	30	62.5	993	1 VIA_TAV	P28931 tomato aspe
32	30	62.5	1088	1 PGDS_RAT	P20786 rattus norv
33	30	62.5	1401	1 RPOC_VIBCH	O9kv29 vibrio chol

34	30	62.5	1498	1 Y1A9_CLOAB	Q04351 clostridium
35	30	62.5	3099	1 POLG_PEMVM	O56075 p genome po
36	29	60.4	126	1 YK50_ARCFU	O38229 archaeoglob
37	29	60.4	154	1 PSB4_PIG	Q29384 sus scrofa
38	29	60.4	218	1 PURQ_METH	O26270 methanobact
39	29	60.4	223	1 PURQ_RHIME	Q92p11 rhizobium m
40	29	60.4	230	1 PURQ_METJA	Q59042 methanococc
41	29	60.4	242	1 PSB4_XENLA	P28024 xenopus lae
42	29	60.4	263	1 MPLA_LOLPR	P14946 lolium pere
43	29	60.4	264	1 PSB4_HUMAN	P28070 homo sapien
44	29	60.4	264	1 PSB4_MOUSE	P99026 mus musculu
45	29	60.4	271	1 POBR_ACICA	P43992 acinetobact
46	29	60.4	298	1 MIAA_AGR5	P38436 agrobacteri
47	29	60.4	354	1 VGLI_VZVD	P09258 varicella-z
48	29	60.4	400	1 ARRB_CALVI	P51487 calliphora
49	29	60.4	400	1 NUCM_PROWI	Q37619 prototheca
50	29	60.4	421	1 ACDM_HUMAN	P11310 homo sapien
51	29	60.4	421	1 ACDM_PIG	P41367 sus scrofa
52	29	60.4	429	1 GSAB_BACSU	P71084 bacillus su
53	29	60.4	432	1 PURA_YEAST	P80210 saccharomyc
54	29	60.4	455	1 HXK2_SCHPO	P50521 schizosacch
55	29	60.4	529	1 ENP3_HUMAN	O75355 homo sapien
56	29	60.4	670	1 OATP_RAT	P46720 rattus norv
57	29	60.4	691	1 OAT6_MOUSE	Q9y616 homo sapien
58	29	60.4	827	1 PLSB_MOUSE	Q61586 mus musculu
59	29	60.4	828	1 PLSB_RAT	P97564 rattus norv
60	29	60.4	831	1 SYFB_MYCTU	P94985 mycobacteri
61	29	60.4	1097	1 RNT1_FUGRU	Q98tr3 fugu rubrip
62	29	60.4	1113	1 RNT1_MOUSE	Q98pu0 mus musculu
63	29	60.4	1129	1 RNT1_HUMAN	Q92900 homo sapien
64	29	60.4	1377	1 RHSA_ECOLI	P16916 escherichia
65	29	60.4	1397	1 RHSC_ECOLI	P16918 escherichia
66	29	60.4	1411	1 RHSE_ECOLI	P15917 escherichia
67	29	60.4	3828	1 TRX_DROVI	Q27442 drosophila
68	28	58.3	119	1 YHHM_ECOLI	P37615 escherichia
69	28	58.3	207	1 ENO_CAMPE	O30885 campylobact
70	28	58.3	223	1 PURQ_PYRAB	Q9uxw5 pyrococcus
71	28	58.3	224	1 PURQ_ANASP	O8yu79 anabaena sp
72	28	58.3	255	1 PRDC_HELPJ	O92mn9 helicobacte
73	28	58.3	255	1 PRDC_HELPY	O08912 helicobacte
74	28	58.3	279	1 ATND_RAT	Q63377 rattus norv
75	28	58.3	286	1 3HAO_RAT	P46953 rattus norv

ALIGNMENTS

RESULT 1
A10C_HUMAN STANDARD; PRT; 1499 AA.
AC O60312; Q96914;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)
DE (Aminophospholipid translocase VC).
GN ATP10C OR ATPVC OR KIAA0566.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2125279; PubMed=11326269;
RA Oshimura M.;
RA "A novel maternally expressed gene, ATP10C, encodes a putative
RT aminophospholipid translocase associated with Angelman syndrome.";
RL Nat. Genet. 28:19-20(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313119; PubMed=11353404;
RA Herzing L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;

DR InterPro: IPR005481; CPase_L_N.
 DR InterPro: IPR004362; MGS_Like.
 DR Pfam: PF00289; CPase_L_Chain; 2.
 DR Pfam: PF02786; CPase_L_D2; 2.
 DR Pfam: PF02787; CPase_L_D3; 1.
 DR Pfam: PF02142; MGS; 1.
 DR PRINTS: PR00098; CPASE.
 DR PROSITE: PS00866; CPASE_1; 2.
 DR PROSITE: PS00867; CPASE_2; 2.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 KW ATP-binding; Manganese; Complete proteome.
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.
 FT REPEAT 1 546
 FT REPEAT 547 1058
 FT NP_BIND 153 210 ATP (POTENTIAL).
 FT NP_BIND 302 352 ATP (POTENTIAL).
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;
 Query Match 72.98; Score 35; DB 1; Length 1058;
 Best Local Similarity 60.08; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXSYS 11
 Db 190 EIVPGLNYS 199
 RESULT 3
 YLIK_TYDVA STANDARD; PRT; 102 AA.
 AC P31619;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE Hypothetical 11.2 kDa protein.
 GN VI.
 OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=31599;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92188538; PubMed=1546458;
 RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;
 RT "The nucleotide sequence of the infectious cloned DNA component of
 tobacco yellow dwarf virus reveals features of geminiviruses
 infecting monocotyledonous plants.";
 RL Virology 187:633-642(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81103; AAA47947.1; -
 DR PIR; A42452; A42452.
 DR InterPro: IPR002621; Gemini_mov.
 DR Pfam: PF01708; Gemini_mov; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;
 Query Match 70.8%; Score 34; DB 1; Length 102;
 Best Local Similarity 60.0%; Pred. No. 2.1;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXSYS 11
 Db 7 QVVPGSINYS 16
 RESULT 4
 GSR2_HUMAN STANDARD; PRT; 478 AA.
 ID GSR2_HUMAN Q9NPM5; Q9NPM1; Q9NPM4; Q9UFI2; Q9BTC6; Q9HAX6;
 AC Q9NPM5; Q9NPM1; Q9NPM4; Q9UFI2; Q9BTC6; Q9HAX6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glioma tumor suppressor candidate region gene 2 protein (p60).
 GN GLTSCR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20175430; PubMed=10708517;
 RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,
 RA Portier B.P., Geki K., Billings S., Ramasamy S., Mohrenweiser H.W.,
 RA Scheithauer B.W., Louis D.N., Jenkins R.B.;
 RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor
 RT region.";
 RL Genomics 64:44-50(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.
 RX MEDLINE=99214318; PubMed=10196275;
 RA Bruni R., Fineschi B., Ogle W.O., Roizman B.;
 RT "A novel cellular protein, p60, interacting with both herpes simplex
 RT virus 1 regulatory proteins ICP22 and ICP0 is modified in a
 RT cell-type-specific manner and is recruited to the nucleus after
 RT infection.";
 RL J. Virol. 73:3810-3817(1999).
 RN [4]
 RP SEQUENCE OF 12-478 FROM N.A.
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 218-477 FROM N.A.
 RC TISSUE=Testis;
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: Interacts with HSV-1 early proteins ICP22 and ICP0.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and
 CC pancreas, moderate levels in placenta, liver, skeletal muscle, and
 CC kidney, and low levels in brain and lung.
 CC -!- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF182076; AAF62873.1; -
 DR EMBL; BC004229; AAH04229.1; -
 DR EMBL; BC006311; AAH06311.1; -
 DR EMBL; BC010095; AAH10095.1; -
 DR EMBL; AF296124; AAG30413.1; -
 DR EMBL; AL359335; CAB94786.1; -

DR	EMBL; X52154; CAA36405.1; -
DR	PIR; S09988; YKLMSI.
DR	HIV; X52154; REVSCPZ.
DR	InterPro; IPR000625; REV_protein.
DR	Pfam; PF00424; REV; 1.
DR	Transcription factor regulation; AIDS; Phosphorylation; Nuclear protein.
DR	SEQUENCE 124 AA; 13701 MW; F5977D1BDF65A7B2 CRC64;
DR	SK

VARIANT	75	75
VARIANT	129	129
T -> A (IN STRAIN	129	129
K > R (IN STRAIN	75	75

VARIANT	75	75
VARIANT	129	129
T -> A (IN STRAIN	129	129
K > R (IN STRAIN	75	75

FT VARIANT 237 237 A -> T (IN STRAIN 1479).
 FT VARIANT 322 322 S -> N (IN STRAIN 1479).
 FT VARIANT 326 326 A -> V (IN STRAIN 1479).
 FT VARIANT 328 328 A -> S (IN STRAIN 1479).
 SQ SEQUENCE 427 AA; 44967 MW; 0882201AEE9254B9 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 427;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 DB 103 QVVPNGSGYS 112

RESULT 7

AROA_VIBCH STANDARD; PRT; 426 AA.
 ID AROA_VIBCH
 AC O9KR80;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-phosphorylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 DE enolpyruvylshikimate-3-phosphate synthase (EPSP synthase) (EPSPS).
 GN AROA OR VC1732.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EI Tor 116961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae";
 RT Nature 406:477-483(2000).
 CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate = phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -|- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway; sixth step.
 CC -|- SUBUNIT: Monomer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE004251; AAF94882.1; -
 DR TIGR; VC1732; -
 DR InterPro; IPR001986; EPSP_synthase.
 DR Pfam; PF00275; EPSP_synthase; 1.
 DR ProDom; PD001867; EPSP_synthase; 1.
 DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
 DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
 KW Aromatic amino acid biosynthesis; Transferrase; Complete proteome.
 SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match 66.7%; Score 32; DB 1; Length 426;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 10

Db 223 EFVIPAGQSY 232
 I | | | | | | |

RESULT 8
 BCNS_CLOPE
 ID BCNS_CLOPE STANDARD; PRT; 890 AA.
 AC P08696;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Bacteriocin BCN5.
 GN BCN.
 OS Clostridium perfringens.
 OG Plasmid pIP404.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=88336297; PubMed=2901768;
 RA Garnier T., Cole S.T.;
 RT "Complete nucleotide sequence and genetic organization of the bacteriocinogenic plasmid, pIP404, from Clostridium perfringens.";
 RL Plasmid 19:134-150(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=87057020; PubMed=2877971;
 RA Garnier T., Cole S.T.;
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium perfringens and molecular genetic analysis of the bacteriocin-encoding gene";
 RL J. Bacteriol. 168:1189-1196(1986).
 RN [3]
 RP SEQUENCE OF 1-14 FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=89039249; PubMed=2460717;
 RA Garnier T., Cole S.T.;
 RT "Studies of UV-inducible promoters from Clostridium perfringens in vivo and in vitro";
 RL Mol. Microbiol. 2:607-614(1988).
 CC -|- FUNCTION: MAY FUNCTION AS AN IONOPHORE.
 CC -|- INDUCTION: BY UV IRRADIATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14481; AAA98248.1; -
 DR EMBL; M32882; AAA98249.1; -
 DR PIR; A30481; A30481.
 DR InterPro; IPR003646; SH3_bac.
 DR SMART; SM00287; SH3b; 3.
 KW Antibiotic; Bacteriocin; Plasmid.
 FT DOMAIN 815 869 HYDROPHOBIC.
 SQ SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;

Query Match 66.7%; Score 32; DB 1; Length 890;
 Best Local Similarity 66.7%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 10

Db 170 EVVPGGFTY 178
 I | | | | | | |

RESULT 9

```

ZEP1_HUMAN STANDARD; PRT; 2717 AA.
ID AC
1 ZEP1_HUMAN STANDARD; PRT; 2717 AA.
2 P15822;
3
4 01-APR-1990 (Rel. 14, Created)
5 01-APR-1990 (Rel. 14, Last sequence update)
6 16-OCT-2001 (Rel. 40, Last annotation update)
7 Zinc finger protein 40 (human immunodeficiency virus type I enhancer-
8 binding protein 1) (HIV-EPI) (Major histocompatibility complex binding
9 protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
10 (PRDII-BF1).
11
12 DE (PRDII-BF1).
13
14 GN HIVP1 OR ZNF40.
15
16 OS Homo sapiens (Human).
17
18 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
19 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
20 NCBI_TaxID=9606;
21
22 [1]
23 SEQUENCE FROM N.A.
24 MEDLINE=90169514; PubMed=2106471;
25
26 RA Fan C.M., Maniatis T.;
27
28 RT "A DNA-binding protein containing two widely separated zinc finger
29 motifs that recognize the same DNA sequence.";
30 Genes Dev. 4:29-42(1990).
31
32 [2]
33 SEQUENCE BY NMR OF 2113-2142.
34 MEDLINE=91064333; PubMed=2248949;
35
36 RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
37 Gronenborn A.M.;
38
39 RT "High-resolution three-dimensional structure of a single zinc finger
40 from a human enhancer binding protein in solution.";
41 Biochemistry 29:9324-9334(1990).
42
43 [3]
44 STRUCTURE BY NMR OF 2087-2142.
45 MEDLINE=92232684; PubMed=1567844;
46
47 RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,
48 Gronenborn A.M.;
49
50 RT "High-resolution solution structure of the double Cys2His2 zinc
51 finger from the human enhancer binding protein MBP-1.";
52 Biochemistry 31:3907-3917(1992).
53
54 -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
55 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
56 NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
57
58 IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
59 OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
60 MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
61 IN T-CELL ACTIVATION.
62
63 -1- SUBCELLULAR LOCATION: Nuclear.
64
65 -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
66
67 -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
68 SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
69 ZINC-FINGER IN-BETWEEN.
70
71 -1- SIMILARITY: STRONG, TO HIVP2.
72
73 -----
74 This SWISS-PROT entry is copyright. It is produced through a collaboration
75 between the Swiss Institute of Bioinformatics and the EMBL outstation -
76 the European Bioinformatics Institute. There are no restrictions on its
77 use by non-profit institutions as long as its content is in no way
78 modified and this statement is not removed. Usage by and for commercial
79 entities requires a license agreement (See http://www.isb-sib.ch/announce/
80 or send an email to license@isb-sib.ch).
81
82 -----
83 EMBL; X51435; CAA35798.1; -
84 PIR; A34203; A34203.
85 PDB; 3ZNF; 15-JAN-92.
86 PDB; 4ZNF; 15-JAN-92.
87 PDB; 1BBO; 31-OCT-93.
88 TRANSFAC; T00497; -
89 Genew; HGNC:4920; HIVP1.
90 MIM; 194540; -
91 InterPro; IPR000822; znf_C2H2.
92 Pfan; PF00096; zf_C2H2; 5.
93 PRINTS; PR00048; ZINC_FINGER.
94 SMART; SM00355; znf_C2H2; 4.
95 PROSITE; PS00028; ZINC_FINGER_C2H2.1; 4.
96

```

16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucyl-tRNA synthetase (EC 6.1.1.4) (Leucine--tRNA ligase) (LeuRS).
 GN LEUS OR LL0816.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis sp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
 CC diphosphate + L-leucyl-tRNA(Leu).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE006315; AAK04914.1; -
 DR InterPro: IPR002302; Leu-tRNA-synthla.
 DR InterPro: IPR002300; tRNA-synt_1a.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00985; TRNASYNTHLEU.
 DR TIGRFAMS; TIGR00396; leuS.bact; 1.
 DR PROSITE; PS00178; AA-TRNA-LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 40 51 "HIGH" REGION.
 FT SITE 609 613 "KMSKS" REGION.
 FT BINDING 612 612 ATP (BY SIMILARITY).
 SQ SEQUENCE 829 AA; 93791 MW; F0F0A43014F6C389 CRC64;
 Query Match 64.6%; Score 31; DB 1; Length 829;
 Best Local Similarity 66.7%; Pred. No. 80;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXS 9
 Db 168 EEVLPDGT 176
 RESULT 14
 ID SULH_SCHPO STANDARD; PRT; 877 AA.
 AC 074377;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable sulfate permease C3H7.02.
 GN SPBC3H7.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weijens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs N., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas R.R., Rochet M., Gallard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
 CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL031261; CAA20298.1; -
 DR InterPro: IPR002645; STAS.
 DR InterPro: IPR001902; Sulfate_transp.
 DR Pfam; PF00916; Sulfate_transp; 1.
 DR Pfam; PF01740; STAS; 1.
 DR TIGRFAMS; TIGR00815; sulp; 1.
 DR PROSITE; PS01130; SLC26A; 1.
 DR PROSITE; PS0801; STAS; 1.
 KW Transport; Transmembrane.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 186 206 POTENTIAL.
 FT TRANSMEM 221 241 POTENTIAL.
 FT TRANSMEM 243 263 POTENTIAL.
 FT TRANSMEM 292 312 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 384 404 POTENTIAL.
 FT TRANSMEM 424 444 POTENTIAL.
 FT TRANSMEM 461 481 POTENTIAL.
 FT TRANSMEM 484 504 POTENTIAL.
 FT TRANSMEM 518 538 POTENTIAL.
 FT TRANSMEM 543 563 POTENTIAL.
 FT DOMAIN 594 747 STAS.
 SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;
 Query Match 64.6%; Score 31; DB 1; Length 877;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGXSYS 11
 Db 148 VVPGMSYA 156
 RESULT 15

SYL_TETH
ID SYL_TETH STANDARD; PRT; 1081 AA.
AC P36422;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)
DE (Ilers).
OS ILSA OR CUPC.
GN Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymena; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RN MEDLINE=92165815; PubMed=13711507;
RX Crank C., Martindale D.W.;
RT "Isoleucyl-tRNA synthetase from the ciliated protozoan Tetrahymena
thermophila. DNA sequence, gene regulation, and leucine zipper
motifs."
RL J. Biol. Chem. 267:4592-4599(1992).
CC -!- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) -> AMP +
diphosphate + L-isoleucyl-tRNA(Ile).
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M30942; AAA30122.1; -
DR PIR; A42399; A42399.
DR HSP; P56690; 11LE.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_1le.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRFAMs; TIGR00392; 1leS; 1.
DR PROSITE; PS00178; AA.TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 53 63 "HIGH" REGION.
FT SITE 607 611 "KMSKS" REGION.
FT BINDING 610 610 ATP (BY SIMILARITY).
SQ SEQUENCE 1081 AA; 124850 MW; D4D72616AFB8C795 CRC64;
Query Match 64.6%; Score 31; DB 1; Length 1081;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 EWPVPGXSY 10
Db :|:|:|
796 KVPKGSY 804
RESULT 16
TKNK_HUMAN
ID TKNK_HUMAN STANDARD; PRT; 121 AA.
AC Q9UHF0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K) {ZNEUROK1}.
GN TAC3.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.

RA Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,
RA O'Hara P.;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20322570; PubMed=10866201;
RA Page N.M., Woods R.J., Gardiner S.M., Lomthiasong K., Gladwell R.T.,
RA Butlin D.J., Manyonda I.T., Lowry P.J.;
RT "Excessive placental neurokinin B secretion during the third trimester
causes pre-eclampsia."
RL Nature 405:797-800(2000).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
MUSCLES (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: In pregnancy, the expression of NKB is
confined to the outer syncytiotrophoblast of the placenta,
significant concentrations of NKB can be detected in plasma as
early as week 9, and plasma concentrations of NKB are grossly
elevated in pregnancy-induced hypertension and pre-eclampsia.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF186112; AAF01430.1; -
DR EMBL; AF216586; AAF76980.1; -
DR Genew; HGNC:11521; TAC3.
DR MIM; 162330; -
DR InterPro; IPR003635; Neurokinin.
DR InterPro; IPR002040; Tachykinin.
DR ProDom; PD020370; Neurokinin; 1.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
KW Amidation; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT PROPEP 17 78 BY SIMILARITY.
FT PEPTIDE 81 90 NEUROKININ B.
FT PROPEP 94 121 BY SIMILARITY.
FT MOD_RES 90 90 AMIDATION (G-91 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
SQ SEQUENCE 121 AA; 13438 MW; 14C9AFE2EE9EDECA CRC64;
Query Match 62.5%; Score 30; DB 1; Length 121;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXG 7
Db :|:|:|
28 EEVVPXG 34
RESULT 17
PURQ_PYRHO
ID PURQ_PYRHO STANDARD; PRT; 223 AA.
AC OS9619;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosylformylglycinamide synthase I (EC 6.3.5.3) (FCAM
synthase I)
DE synthase I)
GN PURQ OR PH1955.
OC Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: Z95151; CAB08429.1; -;
 DR EMBL: AL583924; CAC31174.1; -;
 DR Leproma; ML2219; -;
 DR InterPro; IPR000991; GATase_1.
 DR InterPro; IPR002818; ThiJ.
 DR Pfam; PF01965; ThiJ; 1.
 DR PROSITE; PS00442; GATASE_TYPE_I; 1.
 KW Purine biosynthesis; Ligase; Glutamine amidotransferase;
 KW Complete proteome.
 FT ACT_SITE 87 87 GATASE (BY SIMILARITY).
 SQ SEQUENCE 224 AA; 23818 MW; 717DED17A856E4A6 CRC64;

Query Match 62.5%; Score 30; DB 1; Length 224;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
 III I I I
 Db 47 VVPGGFSY 54

RESULT 20
 PURQ_MYCTU STANDARD; PRT; 224 AA.
 AC P71841;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase I (EC 6.3.5.3) (FGAM
 DE synthase I).
 GN PURQ OR RV0788 OR MT0813 OR WTCY369.32.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultston J.E., Taylor K., Whitehead S., Barrall B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Mollaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
 CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
 CC (formamido)-N(1)-(5-phospho-D-ribosyl)acetamide + L-glutamate.
 CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURQ AND PURL.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: Z80226; CAB02378.1; -;
 DR EMBL: AE006971; AAK45055.1; -;
 DR TIGR; MT0813; -;
 DR TubercuList; RV0788; -;
 DR InterPro; IPR000991; GATase_1.
 DR InterPro; IPR002818; ThiJ.
 DR Pfam; PF01965; ThiJ; 1.
 DR PROSITE; PS00442; GATASE_TYPE_I; 1.
 KW Purine biosynthesis; Ligase; Glutamine amidotransferase;
 KW Complete proteome.
 FT ACT_SITE 87 87 GATASE (BY SIMILARITY).
 SQ SEQUENCE 224 AA; 23633 MW; 46B37215D1171D4E CRC64;

Query Match 62.5%; Score 30; DB 1; Length 224;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXSX 10
 III I I I
 Db 47 VVPGGFSY 54

RESULT 21
 PURQ_CORAM STANDARD; PRT; 225 AA.
 ID PURQ_CORAM
 AC Q9RHX0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase I (EC 6.3.5.3) (FGAM
 DE synthase I).
 DE synthase I).
 GN PURQ.
 OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1697;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 6872;
 RA Yonetani Y., Teshiba S.;
 RT "Sequence analysis of Corynebacterium ammoniagenes purL.";
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
 CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
 CC (formamido)-N(1)-(5-phospho-D-ribosyl)acetamide + L-glutamate.
 CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURQ AND PURL.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AB003162; BAA89450.1; -;
 DR InterPro; IPR000991; GATase_1.
 DR PROSITE; PS00442; GATASE_TYPE_I; 1.

KW Purine biosynthesis; Ligase; Glutamine amidotransferase.
 FT ACT_SITE 87
 SQ SEQUENCE 225 AA; 23803 MW; 11CC6B5E5A53256 CRC64;

Query Match 62.5%; Score 30; DB 1; Length 225;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXS 10
 DB 47 VVPGFSY 54

RESULT 22

SCOA_HELPJ STANDARD; PRT; 232 AA.
 ID SCOA_HELPJ
 AC G9ZLE3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)
 DE (Succinyl-CoA:3-oxoacid CoA-transferase) (OXCT A).
 GN SCOA OR JHP0637.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=65963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + a 3-oxo acid - succinate + a 3-
 CC OXO-acyl-CoA.
 CC -1- SUBUNIT: HETERODIMER OF A SUBUNIT A AND A SUBUNIT B.
 CC -1- SIMILARITY: BELONGS TO THE 3-OXOACID COA-TRANSFERASE SUBUNIT A
 CC FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AE001495; AAD06212.1; -
 DR InterPro; IPR004165; CoA.trans.
 DR InterPro; IPR004163; CoA.transf_1.
 DR Pfam; PF01144; CoA.trans; 1.
 DR PROSITE; PS01273; CoA.TRANSF_1; 1.
 KW Transferase; Complete proteome.
 FT DOMAIN 24 30 COA-BINDING (POTENTIAL).
 SQ SEQUENCE 232 AA; 23525 MW; 1657D27435A0E4AE CRC64;

Query Match 62.5%; Score 30; DB 1; Length 232;
 Best Local Similarity 71.4%; Pred. No. 36;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
 DB 190 EEIVPAG 196

RESULT 23

SCOA_HELPY STANDARD; PRT; 232 AA.
 ID SCOA_HELPY

AC

DT

DT

DT

DE

DE

DE

GN

OS

OC

OC

OX

RN

RP

RX

RA

RA

RA

RA

RA

RA

RA

RA

RA

RA

RA

RT

RT

RL

RN

RP

RX

RA

RA

RA

RA

RA

RA

RT

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

P56006;

01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Succinyl-CoA:3-ketoacid-coenzyme A transferase subunit A (EC 2.8.3.5)

(Succinyl-CoA:3-oxoacid CoA-transferase) (OXCT A).

SCOA OR HP0691.

Helicobacter pylori (Campylobacter pylori).

Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

Helicobacter.

NCBI_TaxID=210;

[1]

SEQUENCE FROM N.A.

STRAIN=26695 / ATCC 700392;

MEDLINE=97394467; PubMed=9252185;

Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,

Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,

McKenney K., Richardson D., Dodson R., Khalak H.G., Glodek A.,

Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,

Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

Venter J.C.;

"The complete genome sequence of the gastric pathogen Helicobacter

pylori.";

Nature 388:539-547(1997).

[2]

SEQUENCE FROM N.A., AND CHARACTERIZATION.

STRAIN=69A;

MEDLINE=97467359; PubMed=9325289;

Corthesy-Theulaz I.E., Bergonzelli G.E., Henry H., Bachmann D.,

Schorderet D.F., Blum A.L., Ornston L.N.;

"Cloning and characterization of Helicobacter pylori succinyl

CoA:acetate CoA-transferase, a novel prokaryotic member of the

CoA-transferase family.";

J. Biol. Chem. 272:25659-25667(1997).

-1- CATALYTIC ACTIVITY: Succinyl-CoA + a 3-oxo acid - succinate + a 3-
 OXO-acyl-CoA.

-1- SUBUNIT: HETERODIMER OF A SUBUNIT A AND A SUBUNIT B.

-1- SIMILARITY: BELONGS TO THE 3-OXOACID COA-TRANSFERASE SUBUNIT A
 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AE000582; AAD07743.1; -

TIGR; HP0691; -

InterPro; IPR004165; CoA.trans.

InterPro; IPR004163; CoA.transf_1.

Pfam; PF01144; CoA.trans; 1.

PROSITE; PS01273; CoA.TRANSF_1; 1.

Transferase; Complete proteome.

DOMAIN 24 30 COA-BINDING (POTENTIAL).

CONFLICT 14 14 A -> T (IN REF. 2).

CONFLICT 108 109 HA -> RP (IN REF. 2).

CONFLICT 116 121 AVYTP -> LITROP (IN REF. 2).

CONFLICT 129 139 QGKSEFNGK -> PQGIQSLTAR (IN REF. 2).

CONFLICT 231 231 T -> A (IN REF. 2).

SEQUENCE 232 AA; 25362 MW; 83CFEE57C0192F3D CRC64;

Query Match 62.5%; Score 30; DB 1; Length 232;

Best Local Similarity 71.4%; Pred. No. 36;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

DB 190 EEIVPAG 196

[1]

[2]

[3]

[4]

[5]

[6]

[7]

[8]

[9]

[10]

[11]

[12]

[13]

[14]

[15]

[16]

[17]

```

RESULT 24
MAG8_HUMAN
ID MAG8_HUMAN STANDARD; PRT; 234 AA.
AC P43361;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma-associated antigen 8 (MAGE-8 antigen).
GN MAGE8 OR MAGE8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012457; PubMed=7927540;
RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL Immunogenetics 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX Timms K.M., Bondeson M.L., Ansari-Lari M.A., Lagerstedt K.,
RA Nelson D.L., Pettersson U., Gibbs R.A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: CONTAINS 1 MAGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10693; AAA68876.1; -
DR Genew; HGNC:6806; MAGEA8.
DR MIM; 300341; -
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Tumor antigen.
FT DOMAIN 112 234
FT DOMAIN 40 43 POLY-SER.
FT SEQUENCE 234 AA; 25197 MW; 058A92EE6003A982 CRC64;
SQ SEQUENCE 234 AA; 25197 MW; 058A92EE6003A982 CRC64;

Query Match 62.5%; Score 30; DB 1; Length 234;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
Db :||| | | | |
170 KEVDPAGHSY 179

RESULT 25
MAG9_HUMAN
ID MAG9_HUMAN STANDARD; PRT; 315 AA.
AC P43362; Q92910;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Melanoma-associated antigen 9 (MAGE-9 antigen).
GN MAGE9 OR MAGE9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Timms K.M., Bondeson M.L., Ansari-Lari M.A., Lagerstedt K.,
RA Nelson D.L., Pettersson U., Gibbs R.A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: CONTAINS 1 MAGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10694; AAA68877.1; -
DR Genew; HGNC:6807; MAGEA9.
DR MIM; 300342; -
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Tumor antigen.
FT DOMAIN 108 307
FT DOMAIN 34 37 POLY-GLU.
FT DOMAIN 87 90 POLY-GLU.
FT SEQUENCE 315 AA; 35088 MW; 7FD2ED10D680D928 CRC64;
SQ SEQUENCE 315 AA; 35088 MW; 7FD2ED10D680D928 CRC64;

Query Match 62.5%; Score 30; DB 1; Length 315;
Best Local Similarity 60.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
Db :||| | | | |
166 KEVDPAGHSY 175

Search completed: June 10, 2003, 13:40:23
Job time : 6.5 secs

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012457; PubMed=7927540;
RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT "Structure, chromosomal localization, and expression of 12 genes of
RT the MAGE family.";
RL Immunogenetics 40:360-369(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX Timms K.M., Bondeson M.L., Ansari-Lari M.A., Lagerstedt K.,
RA Nelson D.L., Pettersson U., Gibbs R.A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA.
CC -!- SIMILARITY: CONTAINS 1 MAGE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10694; AAA68877.1; -
DR Genew; HGNC:6807; MAGEA9.
DR MIM; 300342; -
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS50838; MAGE; 1.
KW Antigen; Multigene family; Tumor antigen.
FT DOMAIN 108 307
FT DOMAIN 34 37 POLY-GLU.
FT DOMAIN 87 90 POLY-GLU.
FT SEQUENCE 315 AA; 35088 MW; 7FD2ED10D680D928 CRC64;
SQ SEQUENCE 315 AA; 35088 MW; 7FD2ED10D680D928 CRC64;

Query Match 62.5%; Score 30; DB 1; Length 315;
Best Local Similarity 60.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSX 10
Db :||| | | | |
166 KEVDPAGHSY 175

Search completed: June 10, 2003, 13:40:23
Job time : 6.5 secs

```


GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-43
Perfect score: 48
Sequence: 1 EEVVPXGXSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	77.1	471	11 Q8R126	Q8R126 mus musculus
2	37	77.1	484	11 Q8VD18	Q8VD18 mus musculus
3	36	75.0	1499	4 Q96914	Q96914 homo sapien
4	36	75.0	3472	1 Q74056	Q74056 cenarchaeum
5	35	72.9	225	10 Q40129	Q40129 lycopersico
6	35	72.9	1063	16 Q8RG86	Q8RG86 fusobacteri
7	34	70.8	143	17 Q8TX62	Q8TX62 methanopyru
8	34	70.8	165	17 Q28330	Q28330 archaeoglob
9	34	70.8	174	10 Q9M3T4	Q9M3T4 betula verr
10	34	70.8	479	4 Q96CS0	Q96CS0 homo sapien
11	34	70.8	541	16 Q98BP5	Q98BP5 rhizobium 1
12	34	70.8	678	12 Q9E1X6	Q9E1X6 cercopithec
13	34	70.8	1028	16 Q8YJ11	Q8YJ11 brucella me
14	34	70.8	1442	17 Q96YH5	Q96YH5 sulfolobus
15	33	68.8	78	6 Q9XST4	Q9XST4 canis famil
16	33	68.8	200	16 Q8XHX3	Q8XHX3 clostridium

RESULT 1

Q8R126
ID Q8R126 PRELIMINARY; PRT; 471 AA.
AC Q8R126: (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 54.5 kDa protein (Fragment).
OS Mus musculus (Mouse).

ALIGNMENTS

17	33	68.8	217	4	O00404	O00404 homo sapien
18	33	68.8	280	5	Q9VMN9	Q9VMN9 drosophila
19	33	68.8	299	4	Q9UEE9	Q9UEE9 homo sapien
20	33	68.8	815	10	Q9SF93	Q9SF93 arabidopsis
21	33	68.8	933	5	Q8SS39	Q8SS39 encephalito
22	32	66.7	96	3	Q9Y855	Q9Y855 kluyveromyc
23	32	66.7	219	17	Q971S2	Q971S2 sulfolobus
24	32	66.7	227	16	Q9RZU8	Q9RZU8 deinococcus
25	32	66.7	290	4	Q96MU1	Q96MU1 homo sapien
26	32	66.7	387	16	Q98FX1	Q98FX1 rhizobium 1
27	32	66.7	387	16	Q92MD6	Q92MD6 rhizobium m
28	32	66.7	425	5	Q9XVK4	Q9XVK4 caenorhabdi
29	32	66.7	556	4	O43733	O43733 homo sapien
30	32	66.7	583	5	Q9BHA5	Q9BHA5 plasmodium
31	32	66.7	583	5	Q9BH83	Q9BH83 plasmodium
32	32	66.7	602	10	Q9LUE0	Q9LUE0 arabidopsis
33	32	66.7	630	2	O87110	O87110 comamonas a
34	32	66.7	670	11	Q01487	Q01487 rattus norv
35	32	66.7	722	12	Q9QTC7	Q9QTC7 marek's dis
36	32	66.7	1735	4	Q9HBL0	Q9HBL0 homo sapien
37	32	66.7	1802	12	Q9J5C2	Q9J5C2 fowlpox vir
38	32	66.7	1902	4	Q14122	Q14122 homo sapien
39	31	64.6	121	2	O33483	O33483 pseudomonas
40	31	64.6	123	16	Q8ZLE6	Q8ZLE6 salmonella
41	31	64.6	123	17	O29517	O29517 archaeoglob
42	31	64.6	193	2	Q8VUA8	Q8VUA8 lactococcus
43	31	64.6	204	11	Q920M7	Q920M7 mus musculu
44	31	64.6	209	16	Q8RE56	Q8RE56 fusobacteri
45	31	64.6	225	11	Q9ET57	Q9ET57 mus musculu
46	31	64.6	266	4	Q96181	Q96181 homo sapien
47	31	64.6	284	16	P74187	P74187 synechocyst
48	31	64.6	301	2	Q9EM57	Q9EM57 salmonella
49	31	64.6	301	2	Q9EM52	Q9EM52 salmonella
50	31	64.6	301	2	Q9ETJ3	Q9ETJ3 salmonella
51	31	64.6	301	2	Q9F7G6	Q9F7G6 salmonella
52	31	64.6	301	2	Q9F7G5	Q9F7G5 salmonella
53	31	64.6	301	2	Q9F7G4	Q9F7G4 salmonella
54	31	64.6	301	2	Q9F7G3	Q9F7G3 salmonella
55	31	64.6	301	2	Q9F7G2	Q9F7G2 salmonella
56	31	64.6	301	2	Q9F7G1	Q9F7G1 salmonella
57	31	64.6	301	2	Q9F7G0	Q9F7G0 salmonella
58	31	64.6	301	2	Q9F7F9	Q9F7F9 salmonella
59	31	64.6	301	2	Q9F7F8	Q9F7F8 salmonella
60	31	64.6	301	2	Q9F7F7	Q9F7F7 salmonella
61	31	64.6	301	2	Q9F7F6	Q9F7F6 salmonella
62	31	64.6	301	2	Q9F7F5	Q9F7F5 salmonella
63	31	64.6	301	2	Q9F7F4	Q9F7F4 salmonella
64	31	64.6	301	2	Q9F7F3	Q9F7F3 salmonella
65	31	64.6	301	2	Q9F7F2	Q9F7F2 salmonella
66	31	64.6	358	9	Q9ZX56	Q9ZX56 mycobacteri
67	31	64.6	361	2	Q9EU31	Q9EU31 salmonella
68	31	64.6	361	2	Q9F7K1	Q9F7K1 salmonella
69	31	64.6	361	2	Q9F7J0	Q9F7J0 salmonella
70	31	64.6	361	2	Q9F7I1	Q9F7I1 salmonella
71	31	64.6	361	2	Q9F7H7	Q9F7H7 salmonella
72	31	64.6	361	2	Q9F7H0	Q9F7H0 salmonella
73	31	64.6	361	16	O8Z5H9	O8Z5H9 salmonella
74	31	64.6	450	16	Q8UB52	Q8UB52 agrobacteri
75	31	64.6	544	16	Q9PQD2	Q9PQD2 ureaplasma

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025810; AAH25810.1; -
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;

Query Match 77.1%; Score 37; DB 11; Length 471;
 Best Local Similarity 60.0%; Pred. NO. 9.3;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 ||:|:|:|:
 Db 226 EVIPAGASYN 235

RESULT 2

Q8VD18 PRELIMINARY; PRT; 484 AA.
 AC Q8VD18
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to glioma tumor suppressor candidate region gene 2.
 GN AW536441.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SALIVARY GLAND;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC017637; AAH17637.1; -
 DR MGD; MGI:2138595; AW536441.
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;

Query Match 77.1%; Score 37; DB 11; Length 484;
 Best Local Similarity 60.0%; Pred. NO. 9.5;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
 ||:|:|:|:
 Db 239 EVIPAGASYN 248

RESULT 3

Q969I4 PRELIMINARY; PRT; 1499 AA.
 AC Q969I4
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative aminophospholipid translocase (Aminophospholipid-transporting ATPase).
 GN ATPase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21225279; PubMed=11326269;
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
 RA Oshimura M.;
 RT "A novel maternally expressed gene, ATP10C, encodes a putative

RT aminophospholipid translocase associated with Angelman syndrome.";
 RL Nat. Genet. 28:19-20(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21313119; PubMed=11353404;
 RA Herzing L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;
 RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
 adjacent to UBE3A and exhibits similar imprinted expression.";
 RL Am. J. Hum. Genet. 68:1501-1505(2001).
 DR EMBL; AB051358; BAB47392.1; -
 DR EMBL; AY029504; AAK33100.1; -
 DR EMBL; AY029487; AAK33100.1; JOINED.
 DR EMBL; AY029488; AAK33100.1; JOINED.
 DR EMBL; AY029489; AAK33100.1; JOINED.
 DR EMBL; AY029490; AAK33100.1; JOINED.
 DR EMBL; AY029491; AAK33100.1; JOINED.
 DR EMBL; AY029492; AAK33100.1; JOINED.
 DR EMBL; AY029493; AAK33100.1; JOINED.
 DR EMBL; AY029494; AAK33100.1; JOINED.
 DR EMBL; AY029495; AAK33100.1; JOINED.
 DR EMBL; AY029496; AAK33100.1; JOINED.
 DR EMBL; AY029497; AAK33100.1; JOINED.
 DR EMBL; AY029498; AAK33100.1; JOINED.
 DR EMBL; AY029499; AAK33100.1; JOINED.
 DR EMBL; AY029500; AAK33100.1; JOINED.
 DR EMBL; AY029501; AAK33100.1; JOINED.
 DR EMBL; AY029502; AAK33100.1; JOINED.
 DR EMBL; AY029503; AAK33100.1; JOINED.
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR001454; Hignase/hydriase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match 75.0%; Score 36; DB 4; Length 1499;
 Best Local Similarity 72.7%; Pred. NO. 54;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
 |||||:|:|:
 Db 469 EEVPRGGSVS 479

RESULT 4

O74056 PRELIMINARY; PRT; 3472 AA.
 AC O74056
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 367.1 kDa protein.
 OS Cenarchaeum symbiosum.
 OC Archaea; Crenarchaeota; Cenarchaeum.
 OX NCBI_TaxID=46770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B;
 RX MEDLINE=98422450; PubMed=9748430;
 RA Schleper C., Delong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
 RA Swanson R.V.;
 RT "Genomic analysis reveals chromosomal variation in natural populations
 of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
 RL J. Bacteriol. 180:5003-5009(1998).
 CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL; AF083072; AAC62699.1; -
 DR InterPro; IPR000515; BPD_transp.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 4.
 DR SMART; SM00320; WD40; 2.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
 KW Hypothetical protein; Repeat; WD repeat.

RT	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
RL	nucleatum strain ATCC 25586. "	72.9%;	Score 35;	DB 16;	Length 1063;		
RL	J. Bacteriol. 184:2005-2018(2002).	60.0%;	Pred. No. 63;				
DR	EMBL; AF010554; AAL94625.1; -	6;	Conservative	2;	Mismatches	2;	Indels
QW	Ligase; Complete proteome.						
SK	SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;						
QY	2 EVVPGXSYS 11						
DB	195 EIVPGLNYS 204						
	1:1111:11						
RESULT 7							
Q8TX62							
ID	Q8TX62	PRELIMINARY;	PRT;	143 AA.			
AC	Q8TX62;						
DT	01-JUN-2002 (TrEMBLrel. 21, Created)						
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)						
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)						
DE	Uncharacterized conserved protein.						
GN	MK0814.						
OS	Methanopyrus kandleri.						
OC	Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;						
OC	Methanopyrus.						
OX	NCBI_TaxID=2320;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN-AV19 / DSM 6324 / JCM 9639;						
RX	MEDLINE=21927647; PubMed=11930014;						
RA	Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,						
RA	Shcherbinina O.V., Shakhova V.V., Belova G.I., Azavind L.,						
RA	Natalie D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,						
RA	Malykh A.G., Koonin E.V., Kozlovskiy S.A.;						
RT	"The complete genome of hyperthermophile Methanopyrus kandleri AV19						
RT	and monophyly of archaeal methanogens.";						
DR	Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).						
DR	EMBL; AF010372; AAM02027.1; -						
KW	Complete proteome.						
QY	SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBBEDD0B CRC64;						
DB	Query Match	70.8%;	Score 34;	DB 17;	Length 143;		
	Best Local Similarity	60.0%;	Pred. No. 12;				
	Matches	6;	Conservative	1;	Mismatches	3;	Indels
QY	1 EEVVPXGXS 10						
DB	75 EELVPOGAGY 84						
	11:1111:11						
RESULT 8							
Q28330							
ID	Q28330	PRELIMINARY;	PRT;	165 AA.			
AC	Q28330;						
DT	01-JAN-1998 (TrEMBLrel. 05, Created)						
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)						
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)						
DE	Hypothetical protein AF1949.						
GN	AF1949.						
OS	Archaeoglobus fulgidus.						
OC	Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;						
OC	Archaeoglobaceae; Archaeoglobus.						
OX	NCBI_TaxID=2234;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=VC-16 / DSM 4304 / ATCC 49558;						
RX	MEDLINE=98049343; PubMed=9389475;						
RA	Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,						
RA	Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,						
RA	Richardson D.L., Kraljic R.J., Graham D.E., Kyrpides N.C.,						
RA	Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,						

RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE000968; AAB89307.1; -;
 DR TIGR; AF1949; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 17588 MW; BBI7054810ADB8 CRC64;
 Query Match 70.8%; Score 34; DB 17; Length 165;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 10
 Db ||| | | | |
 60 EESIPDGASY 69
 RESULT 9
 Q9M3T4 PRELIMINARY; PRT; 174 AA.
 AC Q9M3T4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Acidic endochitinase (EC 3.2.1.14) (Fragment).
 GN PR3A.
 OS Betula verrucosa (White birch) (Betula pendula).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Betula.
 OX NCBI_TaxID=3505;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Hilovaaara-Teijo M., Korhonen M.S., Palva T.E., Kangasjarvi J.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES). CAB66334.1; -;
 DR EMBL; AJ279692; CAB66334.1; -;
 DR HSSP; P23472; 2HVM.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR Pfam; PF00704; Glyco_hydro_18.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase.
 FT NON_TER 1 1
 FT NON_TER 174 174
 SQ SEQUENCE 174 AA; 17936 MW; 834ADCC6B5C76634 CRC64;
 Query Match 70.8%; Score 34; DB 10; Length 174;
 Best Local Similarity 77.8%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGXSYS 11
 Db ||| | | | |
 74 VPPGGGSYS 82
 RESULT 10
 Q96CS0 PRELIMINARY; PRT; 479 AA.
 AC Q96CS0; Q96177;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Glioma tumor suppressor candidate region gene 2 (Fragment).

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 6-479 FROM N.A.
 RC TISSUE=SKIN;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014009; AAH14009.1; -;
 DR EMBL; BC007248; AAH07248.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 479 AA; 54529 MW; 0A6C3A8B476F7E8F CRC64;
 Query Match 70.8%; Score 34; DB 4; Length 479;
 Best Local Similarity 60.0%; Pred. No. 44;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 EVPXGXSYS 11
 Db ||| | | | |
 240 EVAPAGASYN 249
 RESULT 11
 Q98BP5 PRELIMINARY; PRT; 541 AA.
 AC Q98BP5;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Probable DNA ligase.
 GN MLL5481.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003006; BAB51927.1; -;
 DR InterPro; IPR000977; DNA_ligase.
 DR Pfam; PF01068; DNA_ligase; 1.
 DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 DR PROSITE; PS01060; DNA_LIGASE_A3; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 541 AA; 60645 MW; 2EEF705453F28F8 CRC64;
 Query Match 70.8%; Score 34; DB 16; Length 541;
 Best Local Similarity 60.0%; Pred. No. 50;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 10
 Db ||| | | | |
 445 EELVPVGKAY 454
 RESULT 12
 Q9E1X6 PRELIMINARY; PRT; 678 AA.
 ID Q9E1X6

Q9E1X6;
 AC 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 75.9 kDa protein.
 OS Cercopithecine herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=35245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275348; AAC27217.1;
 KW Hypothetical protein.
 SQ SEQUENCE 678 AA; 73850 MW; A17B09E30512FE3C CRC64;
 Query Match 70.8%; Score 34; DB 12; Length 678;
 Best Local Similarity 50.0%; Pred. No. 64;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSX 10
 ||::||| |
 Db 147 EEIIPKGRTRY 156

RESULT 13
 Q8YJ11 PRELIMINARY; PRT; 1028 AA.
 AC Q8YJ11;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP-dependent DNA helicase.
 GN BMEI0275.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujar C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009470; AAU51457.1;
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00490; HELIC_C; 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 1028 AA; 112996 MW; A752B7042572E219 CRC64;

Query Match 70.8%; Score 34; DB 16; Length 1028;
 Best Local Similarity 54.5%; Pred. No. 1e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSXS 11
 ||::||| |
 Db 76 EKIVPPGARYS 86

RESULT 14
 Q96YH5

Q96YH5 PRELIMINARY; PRT; 1442 AA.
 AC Q96YH5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein ST2195.
 GN ST2195.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 Crenarchaeon, Sulfolobus tokodaii strain 7."
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000989; BAB67302.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1442 AA; 156497 MW; D63EC2C35228121F CRC64;
 Query Match 70.8%; Score 34; DB 17; Length 1442;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSX 10
 ||::||| |
 Db 863 EEITPGTGAN 872

RESULT 15
 Q9XST4 PRELIMINARY; PRT; 78 AA.
 AC Q9XST4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE P97 homologous protein (Fragment).
 GN P97.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYROID;
 RX MEDLINE=20422104; PubMed=10964405;
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
 RA Christophe D.;
 RT "A method for the large-scale cloning of nuclear proteins and nuclear
 targeting sequences on a functional basis."
 RL Anal. Biochem. 284:231-239(2000).
 DR EMBL: AJ388531; CAB46833.1;
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8895 MW; B62486313555FBAL CRC64;

Query Match 58.8%; Score 33; DB 6; Length 78;
 Best Local Similarity 54.5%; Pred. No. 10;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXSXS 11
 ||::||| |
 Db 16 EDYVPSGGEYS 26

RESULT 16

Q8HX3
ID Q8HX3 PRELIMINARY; PRT; 200 AA.
AC Q8HX3
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Probable spore cortex-lytic enzyme.
GN CPE2353.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003194; BAB82059.1; -
KW Complete proteome.
SQ SEQUENCE 200 AA; 22163 MW; 3DB0396E5C6B7BE7 CRC64;

Query Match 68.8%; Score 33; DB 16; Length 200;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
Db 142 DVVPDGD SYN 151
:|||||:

RESULT 17
000404
ID 000404 PRELIMINARY; PRT; 217 AA.
AC 000404;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P97 homologous protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=97160586; PubMed=9006920;
RA Nobukuni T., Kobayashi M., Oomori A., Ichinose S., Iwanaga T.,
RA Takahashi I., Hashimoto K., Hattori S., Kaibuchi K., Miyata Y.;
RT "An Alu-linked repetitive sequence corresponding to 280 amino acids is
RT expressed in a novel bovine protein, but not in its human homologue.";
RL J. Biol. Chem. 272:2801-2807(1997).
DR EMBL; D85939; BAA20069.1; -
SQ SEQUENCE 217 AA; 24061 MW; E404BA2E35497828 CRC64;

Query Match 68.8%; Score 33; DB 4; Length 217;
Best Local Similarity 54.3%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
Db 16 EDVVPDGD SYN 26
:|||||:

RESULT 18
Q9VMN9
ID Q9VMN9 PRELIMINARY; PRT; 280 AA.
AC Q9VMN9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CGI4010 protein.
GN CGI4010.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M., Pal
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003610; AAF52273.1; -
DR FlyBase; FBgn0031725; CGI4010.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
SQ SEQUENCE 280 AA; 30697 MW; 3705140822960D43 CRC64;

Query Match 68.8%; Score 33; DB 5; Length 280;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXSYS 11
Db 84 EIVPPGGKYS 93
:|||||:

RESULT 19
Q9UEE9
ID Q9UEE9 PRELIMINARY; PRT; 299 AA.
AC Q9UEE9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BCNT protein (CRANIOFACIAL development protein 1).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98267221; PubMed=9602175;
 RA Takahashi I., Nobukuni T., Ohmori H., Kobayashi M., Tanaka S.,
 RA Ohashima K., Okada T., Masui T., Hashimoto K., Iwashita S.;
 RT "Existence of a bovine LINE repetitive insert that appears in the
 of bovine protein BCNT in ruminant, but not in human, genomes.";
 RL Gene 211:387-394(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB009285; BAA31867.1; -;
 DR EMBL; BC000991; AA00991.1; -;
 SQ SEQUENCE 299 AA; 33593 MW; F4A9E928B669451A CRC64;
 Query Match 68.8%; Score 33; DB 4; Length 299;
 Best Local Similarity 54.5%; Pred. No. 44;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 | | | | |
 DB 16 EDVVPGGGEYS 26

RESULT 20
 Q9SF93
 ID Q9SF93 PRELIMINARY; PRT; 815 AA.
 AC Q9SF93
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Putative phosphatidylinositol-4-phosphate 5-kinase.
 GN F8A24.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Renning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F8A24 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC015985; AAF23244.1; -;
 DR InterPro; IPR003409; MORN.
 DR InterPro; IPR002498; PIP5K.
 DR Pfam; PF02493; MORN; 8.
 DR Pfam; PF01504; PIP5K; 1.
 DR SMART; SM00330; PIP5K; 1.
 KW Kinase.
 SQ SEQUENCE 815 AA; 92091 MW; 3194F009D2C50130 CRC64;
 Query Match 68.8%; Score 33; DB 10; Length 815;
 Best Local Similarity 63.6%; Pred. No. 1.3e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 | | | | |
 DB 49 ELVLPDGEYS 59

RESULT 21
 Q8SS39
 ID Q8SS39 PRELIMINARY; PRT; 933 AA.
 AC Q8SS39;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative ATP-dependent RNA helicase (SKI2 subfamily).
 GN ECU04_0910.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 Encephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 RL EMBL; AL590444; CAD25278.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 933 AA; 106883 MW; CA400EE3A7AFF7CF CRC64;
 Query Match 68.8%; Score 33; DB 5; Length 933;
 Best Local Similarity 54.5%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXSYS 11
 | | | | |
 DB 43 EAVVPVGAAYT 53

RESULT 22
 Q9Y855
 ID Q9Y855 PRELIMINARY; PRT; 96 AA.
 AC Q9Y855;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE SLN1 homologue (Fragment).
 GN SLN1.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2359/152;
 RX MEDLINE=20136098; PubMed=10669872;
 RA Bao W.G., Fukuhara H.;
 RT "The ubiquitin-encoding genes of Kluyveromyces lactis.";
 RL Yeast 16:343-351(2000).
 DR EMBL; AJ243801; CAB50891.1; -;
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00072; Response_reg; 1.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00448; REC; 1.
 KW phosphorylation; Sensory transduction.
 FT NON_TER 1
 SQ SEQUENCE 96 AA; 10971 MW; 5D86B3F852DEE8E7 CRC64;
 Query Match 66.7%; Score 32; DB 3; Length 96;
 Best Local Similarity 54.5%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
| | | | |
Db 6 EEILKNGESYS 16

RESULT 23

Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative ribose 5-phosphate isomerase.
GN ST1302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka K., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7";
RL DNA Res. 8:123-140(2001).
DR ENBL: AP0003985; BAB6348.1;
DR InterPro: IP004788; RpiA.
DR ProDom: PD05813; RpiA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match 66.7%; Score 32; DB 17; Length 219;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EEVVPXGXSYS 10
| | | | |
Db 131 EEVVPGVAY 139

RESULT 24

Q9RZU8 PRELIMINARY; PRT; 227 AA.
AC Q9RZU8
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein DRB0013.
GN DRB0013.
OS Deinococcus radiodurans.
OG Plasmid MP1.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Yamathavan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans RL";
RL Science 286:1571-1577(1999).
DR ENBL: AE001826; AAF12657.1;
DR TIGR: DRB0013;
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 227 AA; 24465 MW; 21C43D2CC42843B8 CRC64;

Query Match 66.7%; Score 32; DB 16; Length 227;
Best Local Similarity 54.5%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXSYS 11
| | | | |
Db 43 ESVLPIGHSFS 53

RESULT 25

Q96MU1 PRELIMINARY; PRT; 290 AA.
AC Q96MU1
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ31891 fis, clone NT2RP7003304, weakly similar to YceA protein
DE homolog ybfQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project";
RL Submitted (OCT-2001) to the ENBL/GenBank/DBJ databases.
DR ENBL: AK056453; BAB71188.1;
DR InterPro: IPR001763; Rhodanese-like.
DR Pfam: PF00581; Rhodanese; 1.
DR Pfam: PF00581; Rhodanese; 1.
SQ SEQUENCE 290 AA; 32972 MW; 41FB7FB0217C1421 CRC64;

Query Match 66.7%; Score 32; DB 4; Length 290;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXS 9
| | | | |
Db 35 EEIVPMGIS 43

Search completed: June 10, 2003, 13:46:35
Job time : 26.7857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 seconds
(without alignments)
46.744 Million cell updates/sec

Title: US-09-909-164-44
Perfect score: 52
Sequence: 1 EEVVPXGXHYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

A_Geneseq_101002:*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	49	94.2	11	23	ABB80546 Hepatitis C virus
2	49	94.2	11	23	ABB80550 Hepatitis C virus
3	49	94.2	11	23	ABB80554 Hepatitis C virus
4	49	94.2	11	23	ABB80555 Hepatitis C virus
5	48	92.3	11	23	ABB80523 Hepatitis C virus
6	48	92.3	11	23	ABB80527 Hepatitis C virus
7	48	92.3	11	23	ABB80531 Hepatitis C virus
8	48	92.3	11	23	ABB80532 Hepatitis C virus
9	48	92.3	11	23	ABB80537 Hepatitis C virus
10	48	92.3	11	23	ABB80541 Hepatitis C virus

11	48	92.3	11	23	ABB80558 Hepatitis C virus
12	48	92.3	11	23	ABB80560 Hepatitis C virus
13	40	76.9	11	23	ABB80544 Hepatitis C virus
14	40	76.9	11	23	ABB80545 Hepatitis C virus
15	40	76.9	11	23	ABB80547 Hepatitis C virus
16	40	76.9	11	23	ABB80548 Hepatitis C virus
17	40	76.9	11	23	ABB80549 Hepatitis C virus
18	40	76.9	11	23	ABB80551 Hepatitis C virus
19	40	76.9	11	23	ABB80552 Hepatitis C virus
20	40	76.9	11	23	ABB80553 Hepatitis C virus
21	40	76.9	11	23	ABB80556 Hepatitis C virus
22	40	76.9	11	23	ABB80557 Hepatitis C virus
23	39	75.0	11	23	ABB80521 Hepatitis C virus
24	39	75.0	11	23	ABB80522 Hepatitis C virus
25	39	75.0	11	23	ABB80524 Hepatitis C virus
26	39	75.0	11	23	ABB80525 Hepatitis C virus
27	39	75.0	11	23	ABB80526 Hepatitis C virus
28	39	75.0	11	23	ABB80528 Hepatitis C virus
29	39	75.0	11	23	ABB80529 Hepatitis C virus
30	39	75.0	11	23	ABB80530 Hepatitis C virus
31	39	75.0	11	23	ABB80533 Hepatitis C virus
32	39	75.0	11	23	ABB80534 Hepatitis C virus
33	39	75.0	11	23	ABB80535 Hepatitis C virus
34	39	75.0	11	23	ABB80536 Hepatitis C virus
35	39	75.0	11	23	ABB80538 Hepatitis C virus
36	39	75.0	11	23	ABB80539 Hepatitis C virus
37	39	75.0	11	23	ABB80540 Hepatitis C virus
38	39	75.0	11	23	ABB80542 Hepatitis C virus
39	39	75.0	11	23	ABB80543 Hepatitis C virus
40	39	75.0	11	23	ABB80559 Hepatitis C virus
41	39	75.0	11	23	ABB80561 Hepatitis C virus
42	39	75.0	11	23	ABB80562 Hepatitis C virus
43	39	75.0	11	23	ABB80563 Hepatitis C virus
44	39	75.0	11	23	ABB80564 Hepatitis C virus
45	39	75.0	11	23	ABB80565 Hepatitis C virus
46	39	75.0	11	23	ABB80566 Hepatitis C virus
47	39	75.0	11	23	ABB80567 Hepatitis C virus
48	39	75.0	11	23	ABB80568 Hepatitis C virus
49	37	71.2	161	21	AA140435 Human OREF199
50	37	71.2	161	21	AA140435 Human OREF protein
51	37	71.2	567	22	AAU39521 Propionibacterium
52	36	69.2	1037	23	ABP39949 Staphylococcus epi
53	34	65.4	126	20	AAW96145 Bovine preprotachy
54	34	65.4	150	22	AAU50013 Propionibacterium
55	34	65.4	379	21	AAV59390 Murine soluble int
56	34	65.4	381	21	AAV59390 Group B Streptococ
57	34	65.4	715	20	AAV27348 Group B Streptococ
58	34	65.4	793	20	AAV27347 Group B Streptococ
59	34	65.4	822	21	AAV91283 Streptococcus agal
60	34	65.4	822	22	AAU00028 Streptococcus agal
61	34	65.4	822	22	AAU00030 Streptococcus poly
62	34	65.4	822	23	ABP28904 Streptococcus poly
63	34	65.4	822	23	ABP29703 Streptococcus pneu
64	34	65.4	823	21	AA12746 Streptococcus pneu
65	34	65.4	824	21	AA12747 Streptococcus pneu
66	34	65.4	825	22	AAU00026 Streptococcus pyog
67	34	65.4	825	22	AAU00029 Streptococcus pyog
68	34	65.4	825	23	ABP26207 Streptococcus poly
69	33	63.5	55	23	AAW50964 yeast dual substra
70	33	63.5	80	22	ABG21296 Novel human diagno
71	33	63.5	205	21	AA12726 Streptococcus pneu
72	33	63.5	205	23	AAU84025 Truncated variant
73	33	63.5	209	22	ABG00168 Novel human diagno
74	33	63.5	334	21	AA12728 Streptococcus pneu
75	33	63.5	334	23	AAU84027 Truncated variant

ALIGNMENTS

RESULT 1
ABB80546
ID ABB80546 standard; peptide: 11 AA.

```

XX AC ABB80546;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #26.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX OS Misc-difference 9
XX FT Modified-site 11 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PN 31-JAN-2002.
XX PN 19-JUL-2001; 2001WO-US23169.
XX PN 21-JUL-2000; 2000US-220101P.
XX PN (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunk TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX CC Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 94.2%; Score 49; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0034;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
Db ||||||| |||
1 EEVVPXGXHYS 11

RESULT 2
ABB80550
ID ABB80550 standard; peptide; 11 AA.
XX AC ABB80550;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #30.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"

```

```

OS Synthetic.
XX Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 9
XX FT Modified-site 11 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PN 31-JAN-2002.
XX PN 19-JUL-2001; 2001WO-US23169.
XX PN 21-JUL-2000; 2000US-220101P.
XX PN (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunk TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX CC Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 94.2%; Score 49; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0034;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
Db ||||||| |||
1 EEVVPXGXHYS 11

RESULT 3
ABB80554
ID ABB80554 standard; peptide; 11 AA.
XX AC ABB80554;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #34.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"

```

FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT XX WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 94.2%; Score 49; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGXHYS 11
RESULT 4
ABB80555
ID ABB80555 standard; peptide; 11 AA.
XX AC ABB80555;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #35.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT XX WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.

XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 94.2%; Score 49; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGXHYS 11
RESULT 5
ABB80523
ID ABB80523 standard; peptide; 11 AA.
XX AC ABB80523;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT XX WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 92.3%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0054;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 DQ ||||| |||
 DB 1 EEVVPXGMHYS 11
 RESULT 6
 ABB80527
 ID ABB80527 standard; peptide; 11 AA.
 AC ABB80527;
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 92.3%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0054;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 DQ ||||| |||
 DB 1 EEVVPXGMHYS 11
 RESULT 7
 ABB80531
 ID ABB80531 standard; peptide; 11 AA.
 AC ABB80531;
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #11.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
 Query Match 92.3%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0054;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
 DB 1 EEVVPXGXHYS 11

RESULT 8
 ABB80532
 ID ABB80532 standard; peptide; 11 AA.
 XX AC ABB80532;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #12.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX PN 31-JAN-2002.
 XX PF 19-JUL-2001; 2001WO-US23169.
 XX PR 21-JUL-2000; 2000US-220101P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 92.3%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0054;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11

DB 1 EEVVPXGXHYS 11
 RESULT 9
 ABB80537
 ID ABB80537 standard; peptide; 11 AA.
 XX AC ABB80537;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #17.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX PN 31-JAN-2002.
 XX PF 19-JUL-2001; 2001WO-US23169.
 XX PR 21-JUL-2000; 2000US-220101P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 92.3%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0054;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
 DB 1 EEVVPXGXHYS 11

RESULT 10
 ABB80541
 ID ABB80541 standard; peptide; 11 AA.
 XX

AC ABB80541;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #21.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 KW Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT
 FT Modified-difference 8
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Llm-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 XX Query Match 92.3%; Score 48; DB 23; Length 11;
 XX Best Local Similarity 90.9%; Pred. No. 0.0054;
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 DB 1 EEVVPXGXHYS 11
 RESULT 11
 ABB80558
 ID ABB80558 standard; peptide; 11 AA.
 XX
 XX ABB80558;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 KW

XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT
 FT Modified-site 8 /note= "Oxymethionine"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Llm-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 XX Query Match 92.3%; Score 48; DB 23; Length 11;
 XX Best Local Similarity 90.9%; Pred. No. 0.0054;
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 DB 1 EEVVPXGXHYS 11
 RESULT 12
 ABB80560
 ID ABB80560 standard; peptide; 11 AA.
 XX
 XX ABB80560;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 KW Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"

PA (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
DR WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
PS
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
PS
XX
XX Query Match 76.9%; Score 40; DB 23; Length 11;
CC Best Local Similarity 81.8%; Pred. No. 0.21;
CC Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGTYS 11
XX
XX RESULT 15
ABB80547
ID ABB80547 standard; peptide; 11 AA.
XX
AC ABB80547;
XX
XX 08-OCT-2002 (first entry)
DT
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #27.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 11 residue 7"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX

PS Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
PS
XX
XX Query Match 76.9%; Score 40; DB 23; Length 11;
CC Best Local Similarity 81.8%; Pred. No. 0.21;
CC Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGTYS 11
XX
XX RESULT 16
ABB80548
ID ABB80548 standard; peptide; 11 AA.
XX
AC ABB80548;
XX
XX 08-OCT-2002 (first entry)
DT
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #28.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 9 residue 7"
FT Modified-site 11 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 76.9%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
||||| ||
Db 1 EEVVPXGTDYS 11

RESULT 17

ABB80549
ID ABB80549 standard; peptide; 11 AA.

XX ABB80549;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #29.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.
31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory
activity useful for treating disorders associated with hepatitis C
virus protease

Claim 17; Page 65; 69pp; English.

The sequence represents a peptide compound of the invention having
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
invention are alpha-ketoamide peptide analogues. The peptides have
virucide activity, and are useful for treating and in the manufacture of
a medicament to treat disorders associated with HCV protease. A
pharmaceutical composition comprising the peptide as an active ingredient
is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 76.9%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
||||| ||
Db 1 EEVVPXGSSYS 11

RESULT 18

ABB80551
ID ABB80551 standard; peptide; 11 AA.

XX ABB80551;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.
31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory
activity useful for treating disorders associated with hepatitis C
virus protease

Claim 17; Page 65; 69pp; English.

The sequence represents a peptide compound of the invention having
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
invention are alpha-ketoamide peptide analogues. The peptides have
virucide activity, and are useful for treating and in the manufacture of
a medicament to treat disorders associated with HCV protease. A
pharmaceutical composition comprising the peptide as an active ingredient
is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 76.9%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
||||| ||
Db 1 EEVVPXGSDYS 11

RESULT 19

ABB80552
ID ABB80552 standard; peptide; 11 AA.

XX AC ABB80552;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #32.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 8 /note= "D-form residue"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX FT WO200208251-A2.
XX PN 31-JAN-2002.
XX PD 19-JUL-2001; 2001WO-US23169.
XX PF 21-JUL-2000; 2000US-220101P.
XX PR (CORV-) CORVAS INT INC.
XX PA Lim-wilby M, Levy OE, Brunck TK;
XX PI WPI; 2002-361643/39.
XX DR Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 76.9%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. NO. 0.21;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGSSYS 11
RESULT 20
ABB80553
ID ABB80553 standard; peptide; 11 AA.
XX AC ABB80553;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #33.
XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW

virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 8 /note= "D-form residue"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX FT WO200208251-A2.
XX PN 31-JAN-2002.
XX PD 19-JUL-2001; 2001WO-US23169.
XX PF 21-JUL-2000; 2000US-220101P.
XX PR (CORV-) CORVAS INT INC.
XX PA Lim-wilby M, Levy OE, Brunck TK;
XX PI WPI; 2002-361643/39.
XX DR Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 76.9%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. NO. 0.21;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EEVVPXGXHYS 11
DB 1 EEVVPXGSSYS 11
RESULT 21
ABB80556
ID ABB80556 standard; peptide; 11 AA.
XX AC ABB80556;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.
XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1

XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 75.0%; Score 39; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.33;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 ||||| ||
 Db 1 EEVVPXGMSYS 11
 RESULT 24
 ABB80522
 ID ABB80522 standard; peptide; 11 AA.
 XX AC ABB80522;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 75.0%; Score 39; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.33;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYS 11
 ||||| ||
 Db 1 EEVVPXGMSYS 11
 RESULT 25
 ABB80524
 ID ABB80524 standard; peptide; 11 AA.
 XX AC ABB80524;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.

CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

xx

SQ Sequence 11 AA;

Query Match 75.0%; Score 39; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.33;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EEVVPXGXHYS 11

|||||||

1 EEVVPXGMDYS 11

Search completed: June 10, 2003, 13:39:11

Job time : 32.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds
(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-44
Perfect score: 52
Sequence: 1 EEVVPXGXHS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCPUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	69.2	1037	4 US-09-134-001C-4794	Sequence 4794, Ap
2	34	65.4	126	2 US-08-879-995A-3	Sequence 3, Appli
3	34	65.4	126	3 US-09-215-096-3	Sequence 3, Appli
4	33	63.5	447	4 US-08-961-083-182	Sequence 182, App
5	33	63.5	600	2 US-08-821-119-19	Sequence 19, Appli
6	33	63.5	600	2 US-08-821-118-2	Sequence 2, Appli
7	33	63.5	763	4 US-08-961-083-66	Sequence 66, Appli
8	33	63.5	796	4 US-08-961-083-56	Sequence 56, Appli
9	32	61.5	738	1 US-08-530-010-3	Sequence 3, Appli
10	32	61.5	738	1 US-08-530-010-5	Sequence 5, Appli
11	32	61.5	738	1 US-08-530-010-7	Sequence 7, Appli
12	32	61.5	738	1 US-08-530-010-9	Sequence 9, Appli
13	32	61.5	738	1 US-08-530-010-11	Sequence 11, Appli
14	32	61.5	738	2 US-08-484-101B-3	Sequence 3, Appli
15	32	61.5	738	2 US-08-484-101B-5	Sequence 5, Appli
16	32	61.5	738	2 US-08-484-101B-7	Sequence 7, Appli
17	32	61.5	738	2 US-08-484-101B-9	Sequence 9, Appli
18	32	61.5	738	2 US-08-484-101B-11	Sequence 11, Appli
19	32	61.5	738	4 US-08-714-524D-3	Sequence 3, Appli
20	32	61.5	738	4 US-08-714-524D-5	Sequence 5, Appli
21	32	61.5	738	4 US-08-714-524D-7	Sequence 7, Appli
22	32	61.5	738	4 US-08-714-524D-9	Sequence 9, Appli
23	32	61.5	738	4 US-08-714-524D-11	Sequence 11, Appli
24	31	59.6	70	4 US-09-134-001C-3950	Sequence 3950, Ap
25	31	59.6	237	4 US-09-134-001C-4185	Sequence 4185, Ap
26	31	59.6	502	4 US-09-342-647-4	Sequence 4, Appli
27	31	59.6	622	2 US-08-459-146-2	Sequence 2, Appli

28	31	59.6	622	2 US-08-459-065-2	Sequence 2, Appli
29	31	59.6	630	4 US-09-342-647-2	Sequence 2, Appli
30	31	59.6	667	4 US-09-342-647-28	Sequence 28, Appli
31	30.5	58.7	268	4 US-09-085-305-8	Sequence 8, Appli
32	30.5	58.7	268	4 US-09-085-305-10	Sequence 10, Appli
33	30.5	58.7	268	4 US-09-085-305-17	Sequence 17, Appli
34	30	57.7	121	4 US-09-152-060-68	Sequence 68, Appli
35	30	57.7	121	4 US-09-152-060-85	Sequence 85, Appli
36	30	57.7	122	2 US-08-879-995A-1	Sequence 1, Appli
37	30	57.7	122	3 US-09-215-096-1	Sequence 1, Appli
38	30	57.7	241	3 US-08-834-776A-2	Sequence 2, Appli
39	30	57.7	328	4 US-09-180-827-10	Sequence 10, Appli
40	30	57.7	386	4 US-09-086-483A-2	Sequence 2, Appli
41	30	57.7	386	4 US-09-130-491-6	Sequence 6, Appli
42	30	57.7	615	2 US-08-663-566A-17	Sequence 17, Appli
43	30	57.7	615	2 US-08-023-610-17	Sequence 17, Appli
44	30	57.7	615	2 US-08-288-065A-17	Sequence 17, Appli
45	30	57.7	615	2 US-08-362-240A-17	Sequence 17, Appli
46	30	57.7	615	5 PCT-US95-10245-17	Sequence 17, Appli
47	30	57.7	747	4 US-09-724-864-36	Sequence 36, Appli
48	30	57.7	1081	2 US-08-843-530B-18	Sequence 18, Appli
49	30	57.7	1117	2 US-08-843-530B-33	Sequence 33, Appli
50	30	57.7	1140	4 US-09-220-081-2	Sequence 2, Appli
51	30	57.7	1140	4 US-09-677-575-2	Sequence 2, Appli
52	29.5	56.7	501	2 US-08-408-095-31	Sequence 31, Appli
53	29	55.8	100	2 US-09-047-125-27	Sequence 27, Appli
54	29	55.8	100	3 US-07-736-335E-27	Sequence 27, Appli
55	29	55.8	261	2 US-08-879-561-1	Sequence 1, Appli
56	29	55.8	289	4 US-09-395-674B-2	Sequence 2, Appli
57	29	55.8	716	2 US-08-766-982-1	Sequence 1, Appli
58	29	55.8	716	4 US-09-296-219-1	Sequence 1, Appli
59	29	55.8	1417	2 US-08-559-303B-78	Sequence 78, Appli
60	29	55.8	1417	3 US-08-781-891-78	Sequence 78, Appli
61	29	55.8	1417	4 US-09-175-828-78	Sequence 78, Appli
62	29	55.8	2109	4 US-08-646-695-6	Sequence 6, Appli
63	29	55.8	2109	5 PCT-US96-06053-6	Sequence 6, Appli
64	28	53.8	46	5 PCT-US95-06266-112	Sequence 112, App
65	28	53.8	59	4 US-08-963-851-14	Sequence 14, Appli
66	28	53.8	102	1 US-08-469-427A-2	Sequence 2, Appli
67	28	53.8	102	2 US-08-609-443B-2	Sequence 2, Appli
68	28	53.8	102	2 US-08-589-063C-2	Sequence 2, Appli
69	28	53.8	102	4 US-08-851-896-2	Sequence 2, Appli
70	28	53.8	133	1 US-08-469-427A-9	Sequence 9, Appli
71	28	53.8	133	2 US-08-609-443B-9	Sequence 9, Appli
72	28	53.8	133	2 US-08-569-063C-9	Sequence 9, Appli
73	28	53.8	133	4 US-08-851-896-9	Sequence 9, Appli
74	28	53.8	150	2 US-08-801-105-2	Sequence 2, Appli
75	28	53.8	188	1 US-08-469-427A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-4794
; Sequence 4794, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4794
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4794

Query Match 69.2%; Score 36; DB 4; Length 1037;
 Best Local Similarity 63.6%; Pred. No. 37;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHS 11
 :||| | |||
 Db 199 KEVVSNGLHYS 209

RESULT 2

US-08-879-995A-3
 ; Sequence 3, Application US/08879995A
 ; Patent No. 5985606
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/879,995A
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0326 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 126 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 163590
 ; US-08-879-995A-3

Query Match 65.4%; Score 34; DB 2; Length 126;
 Best Local Similarity 66.7%; Pred. No. 9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9
 :||| | |||
 Db 28 EQVVPGGGH 36

RESULT 3

US-09-215-096-3
 ; Sequence 3, Application US/09215096
 ; Patent No. 6008194
 ; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/215,096
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/879,995
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0326 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 126 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 163590
 ; US-09-215-096-3
 ; Query Match 65.4%; Score 34; DB 3; Length 126;
 ; Best Local Similarity 66.7%; Pred. No. 9;
 ; Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 ; QY 1 EEVVPXGXH 9
 ; :||| | |||
 ; Db 28 EQVVPGGGH 36
 ; RESULT 4
 ; US-08-961-083-182
 ; ; Sequence 182, Application US/08961083
 ; ; Patent No. 6159469
 ; ; GENERAL INFORMATION:
 ; ; APPLICANT: Choi et. al.
 ; ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 ; ; NUMBER OF SEQUENCES: 452
 ; ; CORRESPONDENCE ADDRESS:
 ; ; ADDRESSEE: Human Genome Sciences, Inc.
 ; ; STREET: 9410 Key West Avenue
 ; ; CITY: Rockville
 ; ; STATE: Maryland
 ; ; COUNTRY: USA
 ; ; ZIP: 20850
 ; ; COMPUTER READABLE FORM:
 ; ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; ; COMPUTER: HP Vectra 486/33
 ; ; OPERATING SYSTEM: MSDOS version 6.2
 ; ; SOFTWARE: ASCII Text
 ; ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961.083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 182:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-182

Query Match 63.5%; Score 33; DB 4; Length 447;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
; : : : :
Db 178 IVPHGHHY 185

RESULT 5

US-08-821-119-19
; Sequence 19, Application US/08821119
; Patent No. 5821104
; GENERAL INFORMATION:
; APPLICANT: Holm, Kaj Andre
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Halkier, Torben
; APPLICANT: Lehmebeck, Jan
; TITLE OF INVENTION: Tripeptidyl Aminopeptidase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5821104o No. 5821104disk of No. 5821104th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821.119
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4107.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:

; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal
; US-08-821-119-19

Query Match 63.5%; Score 33; DB 2; Length 600;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VVPXGXHY 11
; : : : :
Db 31 VPKGWHYS 38

RESULT 6

US-08-821-118-2
; Sequence 2, Application US/08821118
; Patent No. 5989889
; GENERAL INFORMATION:
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDES
; TITLE OF INVENTION: HAVING TRIPEPTIDE AMINOPEPTIDASE
; TITLE OF INVENTION: ACTIVITY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989889o No. 5989889disk of No. 5989889th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821.118
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4107.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-821-118-2

Query Match 63.5%; Score 33; DB 2; Length 600;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VVPXGXHY 11
; : : : :
Db 31 VPKGWHYS 38

RESULT 7

US-08-961-083-66
; Sequence 66, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-66

Query Match 63.5%; Score 33; DB 4; Length 763;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
:|||
Db 184 IVPGDHY 191

RESULT 8
US-08-961-083-56
; Sequence 56, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-56

Query Match 63.5%; Score 33; DB 4; Length 796;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
:|||
Db 185 IVPGDHY 192

RESULT 9
US-08-530-010-3
; Sequence 3, Application US/08530010
; Patent No. 5689055
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M.
; APPLICANT: Chang, Caren
; APPLICANT: Bleecker, Anthony B.
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,010
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-530-010-3

Query Match 61.5%; Score 32; DB 1; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10

Db 498 VVPTGSHF 505
||| | |

RESULT 10

US-08-530-010-5
; Sequence 5, Application US/08530010
; Patent No. 5689055
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M.
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B.
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,010
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515/RFT
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-530-010-5

Query Match 61.5%; Score 32; DB 1; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | |

RESULT 11

US-08-530-010-7
; Sequence 7, Application US/08530010
; Patent No. 5689055
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M.
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B.
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA

ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,010
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-530-010-7

Query Match 61.5%; Score 32; DB 1; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | |

RESULT 12

US-08-530-010-9
; Sequence 9, Application US/08530010
; Patent No. 5689055
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M.
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B.
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,010
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249

; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-530-010-9

Query Match 61.5%; Score 32; DB 1; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
Db 498 VVPTGSHF 505

RESULT 13

US-08-530-010-11
; Sequence 11, Application US/08530010
; Patent No. 5689055

; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M.
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B.
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,010
; FILING DATE:
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249

; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-530-010-11

Query Match 61.5%; Score 32; DB 1; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
Db 498 VVPTGSHF 505

RESULT 14

US-08-484-101B-3
; Sequence 3, Application US/08484101B

; Patent No. 5824868
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
; TITLE OF INVENTION: ETHYLENE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,101B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/
; FILING DATE: 01-JUL-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515-2/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-101B-3

Query Match 61.5%; Score 32; DB 2; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
Db 498 VVPTGSHF 505

RESULT 15

US-08-484-101B-5
; Sequence 5, Application US/08484101B
; Patent No. 5824868

; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
; TITLE OF INVENTION: ETHYLENE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/484,101B
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/
FILING DATE: 01-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/086,555
FILING DATE: 01-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57515-2/RFT
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 5:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-101B-5

Query Match 61.5%; Score 32; DB 2; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | |
Db 498 VVPTGSHF 505

RESULT 16
US-08-484-101B-7
Sequence 7, Application US/08484101B
Patent No. 5824868
GENERAL INFORMATION:
APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard F. Trecartin
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/
FILING DATE: 01-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/086,555
FILING DATE: 01-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57515-2/RFT
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-101B-7

Query Match 61.5%; Score 32; DB 2; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | |
Db 498 VVPTGSHF 505

RESULT 17
US-08-484-101B-9
Sequence 9, Application US/08484101B
Patent No. 5824868
GENERAL INFORMATION:
APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard F. Trecartin
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/
FILING DATE: 01-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/086,555
FILING DATE: 01-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57515-2/RFT
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-101B-9

Query Match 61.5%; Score 32; DB 2; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | :
Db 498 VVPTGSHF 505

RESULT 18

US-08-484-101B-11
; Sequence 11, Application US/08484101B
; Patent No. 5824868
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
; TITLE OF INVENTION: ETHYLENE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,101B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/
; FILING DATE: 01-JUL-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,555
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57515-2/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-101B-11

Query Match 61.5%; Score 32; DB 2; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | :
Db 498 VVPTGSHF 505

RESULT 19

US-08-714-524D-3
; Sequence 3, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4

; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-08-714-524D-3

Query Match 61.5%; Score 32; DB 4; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | :
Db 498 VVPTGSHF 505

RESULT 20

US-08-714-524D-5
; Sequence 5, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-08-714-524D-5

Query Match 61.5%; Score 32; DB 4; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| | :
Db 498 VVPTGSHF 505

RESULT 21

US-08-714-524D-7
; Sequence 7, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-08-714-524D-7

Query Match 61.5%; Score 32; DB 4; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10

Db 498 VVPTGSHF 505
||| | | |

RESULT 22

US-08-714-524D-9
; Sequence 9, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-08-714-524D-9

Query Match 61.5%; Score 32; DB 4; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXHY 10
||| | | |
Db 498 VVPTGSHF 505

RESULT 23

US-08-714-524D-11
; Sequence 11, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; APPLICANT: Blecker, Anthony B
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-08-714-524D-11

Query Match 61.5%; Score 32; DB 4; Length 738;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXHY 10
||| | | |
Db 498 VVPTGSHF 505

RESULT 24

US-09-134-001C-3950
; Sequence 3950, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3950
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3950

Query Match 59.6%; Score 31; DB 4; Length 70;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VVPXGXHY 11
: | | | | |
Db 36 MPKGFHYS 43

RESULT 25

US-09-134-001C-4185
; Sequence 4185, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4185
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4185

Query Match 59.6%; Score 31; DB 4; Length 237;
Best Local Similarity 43.8%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 2; Indels 6; Gaps 1;

Qy 1 EEVY-----PXGXHY 10
||| | | | |
Db 106 EEIVKEKVPWPPSGDHY 121

Search completed: June 10, 2003, 13:51:36
Job time : 10.6429 secs

; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS

ALIGNMENTS

ALIGNMENTS

; FILE REFERENCE: 8331-9002
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 715
; TYPE: PRT
; ORGANISM: group B streptococcus
US-09-252-088-16

Query Match 65.4%; Score 34; DB 9; Length 715;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| |||
Db 243 VVPHGDHY 250

RESULT 2

US-09-252-088-15
; Sequence 15, Application US/09252088
; Publication No. US20030031682A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clment
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Jose
; APPLICANT: MARTIN, Denis
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 8331-9002
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 793
; TYPE: PRT
; ORGANISM: group B streptococcus
US-09-252-088-15

Query Match 65.4%; Score 34; DB 9; Length 793;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
||| |||
Db 321 VVPHGDHY 328

RESULT 3

US-09-765-272-182
; Sequence 182, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi, et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 182:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 182:
US-09-765-272-182

Query Match 63.5%; Score 33; DB 10; Length 447;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
:|:|:|
Db 178 IVPHGGHY 185

RESULT 4

US-09-769-787-38
; Sequence 38, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 398
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-38

Query Match 63.5%; Score 33; DB 9; Length 484;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
:|:|:|
Db 202 IVPHGGHY 209

RESULT 5

US-09-765-272-66
; Sequence 66, Application US/09765272
; Patent No. US20020061545A1

```
;
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 66:
;
; US-09-765-272-66
;
; Query Match 63.5%; Score 33; DB 10; Length 763;
; Best Local Similarity 62.5%; Pred. No. 2.3e+02;
; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 VVPXGXHY 10
; Db 184 IVPHGDIY 191
;
; RESULT 6
; US-09-765-272-56
; Sequence 56, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
;
; US-09-765-272-56
;
; Query Match 63.5%; Score 33; DB 10; Length 796;
; Best Local Similarity 62.5%; Pred. No. 2.4e+02;
; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 VVPXGXHY 10
; Db 185 IVPHGDIY 192
;
; RESULT 7
; US-09-769-787-194
; Sequence 194, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 194
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
;
; US-09-769-787-194
;
; Query Match 63.5%; Score 33; DB 9; Length 826;
; Best Local Similarity 62.5%; Pred. No. 2.5e+02;
; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 VVPXGXHY 10
; Db 204 IVPHGDIY 211
;
; RESULT 8
; US-09-884-465A-8
; Sequence 8, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
```

; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-8

Query Match 63.5%; Score 33; DB 9; Length 838;
Best Local Similarity 62.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
: || | ||
Db 203 IVPBGDHY 210

RESULT 9
US-09-884-465A-7
; Sequence 7, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 840
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-7

Query Match 63.5%; Score 33; DB 9; Length 840;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
: || | ||
Db 205 IVPBGDHY 212

RESULT 10
US-09-884-465A-10
; Sequence 10, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683

; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 840
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-10

Query Match 63.5%; Score 33; DB 9; Length 840;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
: || | ||
Db 182 IVPBGDHY 189

RESULT 11
US-09-884-465A-6
; Sequence 6, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1039
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-6

Query Match 63.5%; Score 33; DB 9; Length 1039;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
: || | ||
Db 202 IVPBGDHY 209

RESULT 12
US-10-176-847-22
; Sequence 22, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
CANCER
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-176-847-22

Query Match 61.5%; Score 32; DB 9; Length 1463;

Best Local Similarity 45.5%; Pred. No. 7.3e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
|||:
Db 438 EEVIQTSSHYT 448

RESULT 13

US-09-881-752A-292
; Sequence 292, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-292

Query Match 59.6%; Score 31; DB 10; Length 44;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXHYS 11
|||:
Db 34 VVPTGTHAS 42

RESULT 14

US-09-815-242-10416
; Sequence 10416, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10416
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10416

Query Match 59.6%; Score 31; DB 10; Length 387;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGXH 9
|::|||
Db 188 EEIPTGGH 195

RESULT 15

US-09-738-626-4919
; Sequence 4919, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4919
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4919

Query Match 59.6%; Score 31; DB 9; Length 622;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 EEVVPXG--XHY 10
|::|||
Db 46 EEIPIGVPNHY 57

RESULT 16

US-09-820-843A-26
; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118

RESULT 21
US-09-909-062-1
; Sequence 1, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavaliabhan, Viyyor M
; APPLICANT: Lovey, Raymond G
; APPLICANT: Jao, Edwin
; APPLICANT: Bennett, Frank
; APPLICANT: McCormick, Jinping L
; APPLICANT: Pike, Russell E
; APPLICANT: Bogen, Stephane L
; APPLICANT: Liu, Yi-Tsung
; APPLICANT: Arasappan, Ashok
; APPLICANT: Pinto, Patrick A
; APPLICANT: Njoroge, F George
; APPLICANT: Ganguly, Ashit
; TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
; FILE REFERENCE: IN01157K-US
; CURRENT APPLICATION NUMBER: US/09/909,062
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,109
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLTATION
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: norvaline-C(=O)
US-09-909-062-1

Query Match 57.7%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
Db 1 EEVVPXG 7

RESULT 22
US-09-909-062-9
; Sequence 9, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavaliabhan, Viyyor M
; APPLICANT: Lovey, Raymond G
; APPLICANT: Jao, Edwin
; APPLICANT: Bennett, Frank
; APPLICANT: McCormick, Jinping L
; APPLICANT: Pike, Russell E
; APPLICANT: Bogen, Stephane L
; APPLICANT: Liu, Yi-Tsung
; APPLICANT: Arasappan, Ashok
; APPLICANT: Pinto, Patrick A
; APPLICANT: Njoroge, F George
; APPLICANT: Ganguly, Ashit
; TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
; FILE REFERENCE: IN01157K-US
; CURRENT APPLICATION NUMBER: US/09/909,062

; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,109
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLTATION
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: norvaline-C(=O)
US-09-909-062-9

Query Match 57.7%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
Db 1 EEVVPXG 7

RESULT 23
US-09-909-062-130
; Sequence 130, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavaliabhan, Viyyor M
; APPLICANT: Lovey, Raymond G
; APPLICANT: Jao, Edwin
; APPLICANT: Bennett, Frank
; APPLICANT: McCormick, Jinping L
; APPLICANT: Pike, Russell E
; APPLICANT: Bogen, Stephane L
; APPLICANT: Liu, Yi-Tsung
; APPLICANT: Arasappan, Ashok
; APPLICANT: Pinto, Patrick A
; APPLICANT: Njoroge, F George
; APPLICANT: Ganguly, Ashit
; TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
; FILE REFERENCE: IN01157K-US
; CURRENT APPLICATION NUMBER: US/09/909,062
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/220,109
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 130
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLTATION
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (7)..(7)
; OTHER INFORMATION: AMIDATION
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)..(6)

; OTHER INFORMATION: norvaline-(C=O)
US-09-909-062-130

Query Match 57.7%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 1 EEVVPXG 7

RESULT 24

US-09-925-297-664
; Sequence 664, Application US/09925297
; Patent No. US20020081659A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105

; CURRENT APPLICATION NUMBER: US/09/925,297

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05989

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 928

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 664

; LENGTH: 72

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-297-664

Query Match 57.7%; Score 30; DB 10; Length 72;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
: | | | |
Db 5 IMPEGXHY 12

RESULT 25

US-09-852-797-68

; Sequence 68, Application US/09852797

; Patent No. US20020172994A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P2

; CURRENT APPLICATION NUMBER: US/09/852,797

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-797-68

Query Match 57.7%; Score 30; DB 9; Length 121;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

Search completed: June 10, 2003, 14:35:45
Job time : 15.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds
(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-44
Perfect score: 52
Sequence: 1 EEVVPXGXHY 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	67.3	233	2 T02590	DNA binding protei
2	35	67.3	308	2 A72207	ftsh proteinase ac
3	35	67.3	360	2 E69086	cell division prot
4	35	67.3	1028	2 AF3286	ATP-dependent DNA
5	35	67.3	1396	2 S36851	L-shaped tail fibe
6	34	65.4	102	2 A42452	V1 protein - tobac
7	34	65.4	126	2 A25905	tachykinin B precu
8	34	65.4	279	2 C75538	hypothetical prote
9	34	65.4	495	2 T28717	hypothetical prote
10	34	65.4	534	2 A69284	coenzyme F420-quin
11	34	65.4	822	2 T46758	hypothetical 92.4K
12	34	65.4	1057	2 F98992	carbamoyl-phosphat
13	33	63.5	46	2 E97985	hypothetical prote
14	33	63.5	124	1 VKLJ51	trans-regulatory s
15	33	63.5	156	2 D82618	conserved hypoteth
16	33	63.5	252	2 AE2001	hypothetical prote
17	33	63.5	412	2 A48702	2-methyl-branched-
18	33	63.5	460	2 S69046	hypothetical prote
19	33	63.5	510	2 G86430	T518.1 protein - A
20	33	63.5	743	2 S38143	hypothetical prote
21	33	63.5	802	2 C95136	conserved domain p
22	33	63.5	819	2 B95136	conserved domain p
23	33	63.5	828	2 E98004	hypothetical prote
24	33	63.5	839	2 G95115	conserved hypoteth
25	33	63.5	853	2 C97985	hypothetical prote
26	33	63.5	855	2 D98004	histidine Motif-Co
27	33	63.5	1039	2 H95115	conserved hypoteth
28	33	63.5	1039	2 D97985	hypothetical prote
29	32	61.5	162	2 A70939	probable monoxigen

30	32	61.5	225	2	S57810	hypothetical prote
31	32	61.5	286	2	C86169	prohibitin 2 [impo
32	32	61.5	311	2	H69194	GMP synthetase, su
33	32	61.5	322	2	AF3211	hypothetical prote
34	32	61.5	364	2	S47540	fructose-bisphosph
35	32	61.5	369	2	T17267	hypothetical prote
36	32	61.5	374	2	G69119	probable aspartate
37	32	61.5	425	2	T24111	hypothetical prote
38	32	61.5	454	2	G96568	probable non-photo
39	32	61.5	519	2	G84598	probable bzrp tran
40	32	61.5	536	2	AI3544	aldehyde dehydroge
41	32	61.5	627	2	A69663	DNA mismatch repai
42	32	61.5	696	2	A91247	phage transposase
43	32	61.5	701	2	S61239	hypothetical prote
44	32	61.5	716	1	JC5061	macrophage-stimula
45	32	61.5	738	2	A48246	ethylene-response
46	32	61.5	902	2	AH2999	conserved hypoteth
47	32	61.5	903	2	A96284	hypothetical prote
48	32	61.5	905	2	A54654	centromere protein
49	32	61.5	1374	2	T16129	hypothetical prote
50	32	61.5	1378	1	I48751	protein-tyrosine k
51	31.5	60.6	126	2	F69305	conserved hypoteth
52	31	59.6	61	2	T40544	probable centromer
53	31	59.6	61	2	G83958	hypothetical prote
54	31	59.6	125	2	C97134	hypothetical prote
55	31	59.6	212	2	E82061	peptide methionine
56	31	59.6	284	2	G69117	diphthine synthase
57	31	59.6	319	2	S03833	hypothetical prote
58	31	59.6	364	1	ADHUB	fructose-bisphosph
59	31	59.6	379	2	A69516	probable aspartate
60	31	59.6	387	1	XUEC	acetyl-CoA C-acylt
61	31	59.6	387	2	E91225	acetyl-CoA transfe
62	31	59.6	387	2	C86072	acetyl-CoA transfe
63	31	59.6	387	2	AI0914	small (beta) chain
64	31	59.6	426	2	S58132	Sisi protein precu
65	31	59.6	427	2	F64064	tolB protein - Hae
66	31	59.6	441	2	G82253	conserved hypoteth
67	31	59.6	487	2	S65811	finger protein (cl
68	31	59.6	545	2	T08564	hypothetical prote
69	31	59.6	622	2	S15009	hypothetical prote
70	31	59.6	653	2	D82352	iron(III) ABC tran
71	31	59.6	668	2	C96501	hypothetical prote
72	31	59.6	704	2	T29996	hypothetical prote
73	31	59.6	761	2	T33816	hypothetical prote
74	31	59.6	772	2	T02098	probable phosphati
75	31	59.6	991	2	S43891	dna exoribonucleas

ALIGNMENTS

RESULT 1

T02590
DNA binding protein EREBP-2 - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02590

R;Omme-Takagi, M.; Shinshi, H.
Plant Cell 7, 173-182, 1995

A;Title: Ethylene-inducible DNA binding proteins that interact with an ethylene res

A;Reference number: Z14671; MUID:95276459; PMID:7756828

A;Accession: T02590

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-233 <OHM>

A;Cross-references: EMBL:D38126; NID:G790362; PIDN:BAA07324.1; PID:g1208498

A;Experimental source: strain BY4; tissue-type leaf

Query Match 67.3%; Score 35; DB 2; Length 233;
Best Local Similarity 60.0%; Pred. No. 9.7;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10

Db 90 QAVVPKGRHY 99
: ||| | ||

RESULT 2

A72207
ftsH proteinase activity modulator HflK - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: A72207
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: A72207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <ARN>

A:Cross-references: GB:AE001819; GB:AE000512; NID:g4982396; PIDN:AAD36885.1; PID:g498240

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1822

C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 67.3%; Score 35; DB 2; Length 308;

Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10

||| | ||

Db 41 VVPSGIHY 48

||| | ||

RESULT 3

E69086

cell division protein - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999

C:Accession: E69086

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Olui, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: E69086

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-360 <MTH>

A:Cross-references: GB:AE000923; GB:AE000666; NID:g2622766; PIDN:AAB86115.1; PID:g262277

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1642

C:Superfamily: cell division protein MJ0174

Query Match 67.3%; Score 35; DB 2; Length 360;

Best Local Similarity 45.5%; Pred. No. 16;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 11

||:| | ||:

Db 98 EDLVPMSGSHY 108

||:| | ||:

RESULT 4

AF3286

ATP-dependent DNA helicase BMEI0275 [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AF3286

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; i
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella m

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3286

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1028 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51457.1; PID:g17982167; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0275

A:Map position: 1

Query Match 67.3%; Score 35; DB 2; Length 1028;

Best Local Similarity 54.5%; Pred. No. 48;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 11

||:| | ||

Db 76 EKIVPPGARYS 86

||:| | ||

RESULT 5

S36851

L-shaped tail fiber protein - phage T5

N:Alternate names: ltf protein

C:Species: phage T5

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000

C:Accession: S65934; S01984; S36851

R:Kallman, A.V.; Kuishin, V.E.; Shlyapnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.

FEBS Lett. 366, 46-48, 1995

A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.

A:Reference number: S65934; MUID:95309401; PMID:7789514

A:Accession: S65934

A:Molecule type: DNA

A:Residues: 1-1396 <KAL>

A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416

R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.

Nucleic Acids Res. 16, 6230, 1988

A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between earl

A:Reference number: S01984; MUID:88289370; PMID:3267228

A:Accession: S01984

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 934-985, 'A', 987-1396 <KA2>

A:Cross-references: EMBL:X07559

C:Genetics:

A:Gene: ltf

C:Keywords: late protein; tail fiber

Query Match 67.3%; Score 35; DB 2; Length 1396;

Best Local Similarity 50.0%; Pred. No. 67;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHY 11

: | | | |

Db 1360 KTIPAGDHY 1369

: | | | |

RESULT 6

A42452

VI protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A;Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 65.4%; Score 34; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGXHS 11
||| | ||
Db 7 QVVPGGINYS 16

RESULT 7

A25905
tachykinin B precursor - bovine
N;Alternate names: neuromedin K
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A25905
R;Kotani, H.; Hoshimaru, M.; Nawa, H.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 7074-7078, 1986
A;Title: Structure and gene organization of bovine neuromedin K precursor.
A;Reference number: A25905; MUID:86313713; PMID:3462746
A;Accession: A25905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-126 <KOT>
A;Cross-references: GB:M14351; NID:g163587; PIDN:AAA30723.1; PID:g163590
C;Superfamily: neurokinin B precursor
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-126/Product: protachykinin B #status predicted <MAT>

Query Match 65.4%; Score 34; DB 2; Length 126;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9
|:|:| | |
Db 28 EQVVPGGGH 36

RESULT 8

C75538
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C;Accession: C75538
R;Whiter, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-279 <WHI>
A;Cross-references: GB:AE001889; GB:AE000513; NID:g6457944; PIDN:AAF09867.1; PID:g645795
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0271
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0271

Query Match 65.4%; Score 34; DB 2; Length 279;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPGXGXHS 11
| | | | |
Db 100 VPLGRHYS 107

RESULT 9

T28717
hypothetical 92.4K protein - Streptococcus agalactiae
C;Species: Streptococcus agalactiae
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000

hypothetical protein F10D2.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28717
R;Graves, T.; Wohlmann, P.; Gillam, B.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of C. elegans cosmid F10D2.
A;Reference number: Z20515
A;Accession: T28717
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-495 <GRA>
A;Cross-references: EMBL:AF022972; PIDN:AAC48234.1; GSPDB:GN00023; CESP:F10D2.3
A;Experimental source: strain Bristol N2; clone F10D2
C;Genetics:
A;Gene: CESP:F10D2.3
A;Map position: 5
A;Introns: 37/2; 90/1; 113/1; 183/3; 356/3; 380/3; 428/2

Query Match 65.4%; Score 34; DB 2; Length 495;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
|:|:| | |
Db 218 ENIVPTGKHH 227

RESULT 10

A69284
coenzyme F420-quinone oxidoreductase (EC 1.6.5.-) 56K chain - Archaeoglobus fulgidus
N;Alternate names: sarcosine oxidase alpha chain soxa homolog
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 19-May-2000
C;Accession: A69284; S45665
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.;
Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Syk
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing a
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69284
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-534 <KLE>
A;Cross-references: GB:AE001086; GB:AE000782; NID:g2689409; PIDN:AAB90959.1; PID:g
R;Kunow, J.; Linder, D.; Stetter, K.O.; Thauer, R.K.
Eur. J. Biochem. 223, 503-511, 1994
A;Title: F(420)H(2): quinone oxidoreductase from Archaeoglobus fulgidus. Character
A;Reference number: S45665; MUID:94333340; PMID:8055920
A;Accession: S45665
A;Molecule type: protein
A;Residues: 2,'X',4,'X',6-7,'X',9-10,'XX',13-14 <KUN>
A;Note: the authors could not distinguish between glutamate and cysteine for resid
C;Superfamily: glutamate synthase small chain
C;Keywords: oxidoreductase

Query Match 65.4%; Score 34; DB 2; Length 534;
Best Local Similarity 45.5%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
:|:| | | | | |
Db 119 DKVFPFGSHYT 129

RESULT 11

T46758
hypothetical 92.4K protein - Streptococcus agalactiae
C;Species: Streptococcus agalactiae
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000

C:Accession: T46758
 R:Spellerberg, B.; Rodzinski, E.; Martin, S.; Weber-Heynemann, J.; Schnitzler, N.; Luet
 Infect. Immun. 67, 871-878, 1999
 A:Title: Lmb, a protein with similarities to the Lrai adhesin family, mediates attachment
 A:Reference number: 224091; MUID:99115568; PMID:9916102
 A:Accession: T46758
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-822 <SPE>
 A:Cross-references: EMBL:AF062533; NID:g4249622; PIDN:AAD13797.1; PID:g4249624
 A:Experimental source: strain R268
 C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 65.4%; Score 34; DB 2; Length 822;
 Best Local Similarity 75.0%; Pred. No. 62;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VWPXGXHY 10
 ||| | ||
 DB 350 VVPHGDHY 357

RESULT 12
 F89892
 carbamoyl-phosphate synthase large chain [imported] - Staphylococcus aureus (strain N315
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: F89892
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: F89892
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1057 <KUR>
 A:Cross-references: GB:BA000018; PID:g13701002; PIDN:BAB42298.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: pyrAB
 C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin

Query Match 65.4%; Score 34; DB 2; Length 1057;
 Best Local Similarity 60.0%; Pred. No. 81;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHY 11
 ||| | |||
 DB 190 EIVSNGHY 199

RESULT 13
 E97985
 hypothetical protein spr0909 [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: E97985
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: E97985
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-46 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK99713.1; PID:g15458516; GSPDB:GN00174
 C:Genetics:
 A:Gene: spr0909

Query Match 63.5%; Score 33; DB 2; Length 46;
 Best Local Similarity 62.5%; Pred. No. 4.5;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
 :|| | ||
 DB 21 IVPHGGHY 28

RESULT 14
 VKLJSI
 trans-regulatory splicing protein - simian immunodeficiency virus SIVcpz
 N:Alternate names: anti-repression trans-activator; art protein; rev protein; trs
 C:Species: simian immunodeficiency virus SIVcpz
 A:Note: host Pan troglodytes (Chimpanzee)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
 C:Accession: S09988
 R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
 Nature 345, 356-359, 1990
 A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
 A:Reference number: S09983; MUID:90259077; PMID:2188136
 A:Accession: S09988
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-124 <HUE>
 A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36405.1; PID:g763085
 C:Genetics:
 A:Gene: rev; trs; art
 A:Introns: 27/1
 C:Superfamily: AIDS trans-regulatory splicing protein
 C:Keywords: AIDS; immunodeficiency; splicing protein; transcription regulation

Query Match 63.5%; Score 33; DB 1; Length 124;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHY 11
 ||| | ||
 DB 107 ETVPAGGNY 116

RESULT 15
 D82618
 conserved hypothetical protein XF1950 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: D82618
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide S
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: D82618
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-156 <SIN>
 A:Cross-references: GB:AE004014; GB:AE003849; NID:g9107044; PIDN:AAF84752.1; GSPDB:
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Car
 a-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.;
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.;
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, R.C.; Miyaki,
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Sa, R.G.; Santelli, R.V.; Sa
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Si
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.
 A:Reference number: A59328
 A:Contents: annotation

C:Genetics:
A:Gene: XF1950

Query Match 63.5%; Score 33; DB 2; Length 156;
Best Local Similarity 55.6%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9
|||:| | |
Db 119 EEILPQGVH 127

RESULT 16

AE2001
hypothetical protein alr1563 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AE2001
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077929.1; PID:g17135383; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1563

Query Match 63.5%; Score 33; DB 2; Length 252;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
|||:| | |
Db 235 EMIVPAGLHF 244

RESULT 17

AE48702
2-methyl-branched-chain-enoyl-CoA reductase (EC 1.3.1.52) precursor, mitochondrial - pig
C:Species: Ascaris suum (pig roundworm)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 13-Aug-1999
C:Accession: A48702
R:Duran, E.; Komuniecki, R.W.; Komuniecki, P.R.; Wheelock, M.J.; Klingbeil, M.M.; Ma, Y.
J. Biol. Chem. 268, 22391-22396, 1993
A:Title: Characterization of cDNA clones for the 2-methyl branched-chain enoyl-CoA reductase from Ascaris suum.
A:Reference number: A48702; MUID:94042988; PMID:7693666
A:Accession: A48702
A:Molecule type: mRNA
A:Residues: 1-412 <DUR>
A:Cross-references: GB:L22165; NID:g347402; PIDN:AAA16096.1; PID:g347404
A:Experimental source: anaerobic mitochondria of adult stage muscle
A:Note: In the authors translation residues 149-165 are shown to the right of residues 1
C:Superfamily: acyl-CoA dehydrogenase
A:Keywords: flavoprotein; mitochondrion; oxidoreductase
F:1-16/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:17-412/Product: 2-methyl-branched-chain-enoyl-CoA reductase #status experimental <MAT>

Query Match 63.5%; Score 33; DB 2; Length 412;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
|||:| | |
Db 47 EEVIPRAHH 56

RESULT 18

S69046
hypothetical protein YPL139c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 17-Mar-2000
C:Accession: S69046
R:Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.; Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A:Reference number: S69040
A:Accession: S69046
A:Molecule type: DNA
A:Residues: 1-460 <HAL>
A:Cross-references: EMBL:U43703; NID:g1244769; PIDN:AAB68221.1; PID:g1244776; MIPS:
C:Genetics:
A:Gene: SGD:UME1
A:Cross-references: SGD:S0006060; MIPS:YPL139c
A:Map position: 16L
C:Superfamily: Saccharomyces cerevisiae transcription modulator WTM1

Query Match 63.5%; Score 33; DB 2; Length 460;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
|||:| | |
Db 85 IVPLGLHY 92

RESULT 19

G86430
T518.1 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: G86430
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A.
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewa
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mar
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Ta
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86430
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
A:Cross-references: GB:AE005172; NID:g4587512; PIDN:AAD25743.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: hexose phosphate transport protein uhpT

Query Match 63.5%; Score 33; DB 2; Length 510;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
|||:| | |
Db 12 EEVPPGGIHF 21

RESULT 20

S38143
hypothetical protein YBL011w homolog YKR067w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Apr-2002
C:Accession: S38143
R:van Vliet-Reedijk, J.C.; Planta, R.J.
submitted to the Protein Sequence Database, March 1994

A:Reference number: S38130
 A:Accession: S38143
 A:Molecule type: DNA
 A:Residues: 1-743 <VAN>
 A:Cross-references: EMBL:Z28292; NID:g486537; MIPS:YKR0674
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:GPT2
 A:Cross-references: SGD:S0001775
 A:Map position: 11R
 C:Keywords: transmembrane protein

Query Match 63.5%; Score 33; DB 2; Length 743;
 Best Local Similarity 75.0%; Pred. No. 91;
 Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

QY 3 VVPXGXHY 10
 ||| |
 Db 294 VVPCGLHY 301

RESULT 21
 C95136
 conserved domain protein Sp1175 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: C95136
 R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap
 pple, nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95136
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-802 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75284.1; PID:g14972655; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: Sp1175
 C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 63.5%; Score 33; DB 2; Length 802;
 Best Local Similarity 62.5%; Pred. No. 99;
 Matches 5; Conservative 1; Mismatches 1; Indels 2; Gaps 0;

QY 3 VVPXGXHY 10
 ||| |
 Db 191 IVPHGDHY 198

RESULT 22
 B95136
 conserved domain protein Sp1174 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: B95136
 R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap
 pple, nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: B95136
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-819 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75283.1; PID:g14972654; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:

A:Gene: Sp1174
 C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 63.5%; Score 33; DB 2; Length 819;
 Best Local Similarity 62.5%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
 ||| |
 Db 203 IVPHGDHY 210

RESULT 23

E98004

hypothetical protein phtA [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: E98004

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren,
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas,
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: E98004

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-828 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK99865.1; PID:g15458683; GSPDB:GN00174

C:Genetics:

A:Gene: phtA

C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 63.5%; Score 33; DB 2; Length 828;

Best Local Similarity 62.5%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10

||| |

Db 217 IVPHGDHY 224

RESULT 24

G95115

conserved hypothetical protein Sp1003 [imported] - Streptococcus pneumoniae (strain
 C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: G95115

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.;

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap

pple, nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morr

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95115

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-839 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK75120.1; PID:g14972476; GSPDB:GN00164; TIGR

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: Sp1003

C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 63.5%; Score 33; DB 2; Length 839;

Best Local Similarity 62.5%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10

||| |

Db 204 IVPHGDHY 211

RESULT 25

C97985
 hypothetical protein phtD [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: C97985
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: C97985
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-853 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK99711.1; PID:g15458514; GSPDB:GN00174
 C:Genetics:
 A:Gene: phtD
 C:Superfamily: Streptococcus agalactiae hypothetical 92.4K protein

Query Match 63.5%; Score 33; DB 2; Length 853;
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
 :|||
 Db 203 IVPHGHHY 210

Search completed: June 10, 2003, 13:49:17
 Job time : 12.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-44
Perfect score: 52
Sequence: 1 EEWVPGXHYHS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	67.3	1058	1	CARB_FUSNN	Q8r986 fusbacteri
2	35	67.3	1396	1	VITF_BPT5	P13390 bacterioph
3	34	65.4	102	1	Y11K_TYDVA	P31619 tobacco yel
4	34	65.4	126	1	TKNK_BOVIN	P08858 bos taurus
5	34	65.4	1057	1	CARB_STAAM	Q99ur5 staphylococ
6	34	65.4	1057	1	CARB_STAAM	P58940 staphylococ
7	33	63.5	124	1	REV_SIVCZ	P17280 chimpanzee
8	33	63.5	460	1	UMEL_YEAST	Q03010 saccharomyc
9	33	63.5	743	1	YK47_YEAST	P36148 saccharomyc
10	32	61.5	308	1	GAAB_METH	O26806 methanobact
11	32	61.5	363	1	ALFB_SHEEP	P52210 ovine aries
12	32	61.5	627	1	MUTL_BACSU	P49850 bacillus su
13	32	61.5	735	1	ETRL_BRAOL	O49230 brassica ol
14	32	61.5	738	1	ETRL_ARATH	P49333 arabidopsis
15	32	61.5	906	1	CNCC_MOUSE	P49452 mus musculu
16	32	61.5	1374	1	Y056_CAEEL	Q09541 caenorhabdi
17	32	61.5	1378	1	RON_MOUSE	O62190 mus musculu
18	31	59.6	130	1	S205_RAT	P97885 rattus norv
19	31	59.6	212	1	MSRA_VIBCH	Q9kp30 vibrio chol
20	31	59.6	319	1	YHAL_CRYPA	P10941 cryphonectr
21	31	59.6	363	1	ALFB_HUMAN	P05062 homo sapien
22	31	59.6	363	1	ALFB_RABIT	P79226 oryctolagus
23	31	59.6	387	1	THIK_ECOLI	P21151 escherichia
24	31	59.6	426	1	SUSL_YARLI	Q99158 yarrowia li
25	31	59.6	427	1	TOLB_HAEIN	P44677 haemophilus
26	31	59.6	991	1	DHP1_SCHPO	P40848 schizosacch
27	31	59.6	1188	1	KPBA_CAEEL	P34335 caenorhabdi
28	31	59.6	1410	1	PDRB_YEAST	P40550 saccharomyc
29	31	59.6	1499	1	A10C_HUMAN	O60312 homo sapien
30	31	59.6	3099	1	POUG_FEMVN	O56075 p genome po
31	30.5	58.7	268	1	ILIA_BOVIN	P08831 bos taurus
32	30.5	58.7	268	1	ILIA_CAPHI	P79161 capra hircu
33	30.5	58.7	268	1	ILIA_SHEEP	Q28579 ovine aries

34	30	57.7	121	1	TKNK_HUMAN	Q9uhf0 homo sapien
35	30	57.7	152	1	Y16D_BPT4	P22917 bacterioph
36	30	57.7	160	1	Y201_METJA	Q57654 methanococc
37	30	57.7	223	1	COAT_CIV36	Q00686 citrus tris
38	30	57.7	232	1	SCOA_HELPJ	Q2ie3 helicobacte
39	30	57.7	232	1	SCOA_HELPJ	P56006 helicobacte
40	30	57.7	316	1	COQ3_YEAST	P27680 s hexapreny
41	30	57.7	328	1	ROPA_BORPE	P37368 bordetella
42	30	57.7	353	1	T2BA_BACAR	P19887 bacillus an
43	30	57.7	386	1	T10D_HUMAN	Q9ubn6 homo sapien
44	30	57.7	394	1	NUCC_SYNY3	P27724 synchocyst
45	30	57.7	427	1	TRBI_AERPE	Q9y8t5 aeropyrum p
46	30	57.7	430	1	FOLC_BACSU	Q05865 bacillus su
47	30	57.7	513	1	PHSL_DESBA	P13065 desulfovibr
48	30	57.7	541	1	LNT_CHLPN	Q9z7q1 chlamydia p
49	30	57.7	695	1	VNUC_MABVP	P35263 marburg vir
50	30	57.7	960	1	CSEL_YEAST	P33307 saccharomyc
51	29.5	56.7	526	1	NADB_MYCLE	Q49617 mycobacteri
52	29.5	56.7	527	1	NADB_MYCTU	O06595 mycobacteri
53	29.5	56.7	847	1	CD22_HUMAN	P20273 homo sapien
54	29	55.8	223	1	Y864_THEVO	Q97ae6 thermoplas
55	29	55.8	224	1	Y764_THEAC	Q9hk42 thermoplas
56	29	55.8	227	1	ID11_MESAU	O35586 mesocricetu
57	29	55.8	231	1	FLGH_PSEPU	Q52081 pseudomonas
58	29	55.8	238	1	GLO2_HAEIN	P1374 haemophilus
59	29	55.8	242	1	PSB4_XENLA	P28024 xenopus lae
60	29	55.8	263	1	CANS_BOVIN	P13135 bos taurus
61	29	55.8	285	1	IF2A_PYRAE	Q82ty5 pyrobaculum
62	29	55.8	276	1	PRXC_SYNY3	Q55921 synchococc
63	29	55.8	293	1	YD11_METJA	Q58707 methanococc
64	29	55.8	341	1	HYPE_AZOV1	P04095 azotobacter
65	29	55.8	354	1	VGLI_VZVD	P09258 varicella-z
66	29	55.8	360	1	ALF_DROME	P07764 drosophila
67	29	55.8	363	1	ALFB_CHICK	P07341 gallus gall
68	29	55.8	363	1	ALFB_RAT	P00884 rattus norv
69	29	55.8	363	1	ALF_SCHMA	P53442 schistosoma
70	29	55.8	394	1	MNN9_YEAST	P39107 saccharomyc
71	29	55.8	404	1	YJ9N_YEAST	P47177 saccharomyc
72	29	55.8	421	1	ACDM_HUMAN	P11310 homo sapien
73	29	55.8	421	1	ACDM_PIG	P41367 sus scrofa
74	29	55.8	450	1	Y325_HAEIN	P44640 haemophilus
75	29	55.8	471	1	TMLH_NEUCR	Q96ub1 neurospora

ALIGNMENTS

RESULT 1
CARB_FUSNN STANDARD: PRT: 1058 AA.
ID CARB_FUSNN
AC Q8R986;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
DE CARB OR FN0422.
GN Fusobacterium nucleatum (subsp. nucleatum).
OS Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RP [1]
RC SEQUENCE FROM N.A.
RA MEDLINE-21886394; PubMed-11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasieva O., Chou L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyripides N., Overbeek R.;
RA "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586".
RL J. Bacteriol. 184:2005-2018(2002).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP +

CC phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- COPACTOR: Binds three manganese ions (By similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- PATHWAY: Pyrimidine biosynthesis;
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
 CC promotes the hydrolysis of glutamine to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE010554; AAL94625.1; ALT_INIT.
 CC InterPro; IPR005483; CPase_L.
 CC InterPro; IPR005479; CPase_L_D2.
 CC InterPro; IPR005480; CPase_L_D3.
 CC InterPro; IPR005481; CPase_L_N.
 CC InterPro; IPR004362; MGS_like.
 CC Pfam; PF00289; CPase_L_chain; 2.
 CC Pfam; PF02786; CPase_L_D2; 2.
 CC Pfam; PF02787; CPase_L_D3; 1.
 CC Pfam; PF02142; MGS; 1.
 CC PRINTS; PRO0098; CPASE.
 CC PROSITE; PS00866; CPASE_1; 2.
 CC PROSITE; PS00867; CPASE_2; 2.
 CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 CC ATP-binding; Manganese; Complete proteome.
 CC FT DOMAIN 1 401
 CC FT DOMAIN 402 546
 CC FT DOMAIN 547 929
 CC FT DOMAIN 930 1058
 CC FT REPEAT 1 546
 CC FT REPEAT 547 1058
 CC FT NP_BIND 153 210
 CC FT NP_BIND 302 352
 CC FT METAL 284 284
 CC FT METAL 298 298
 CC FT METAL 300 300
 CC FT METAL 820 820
 CC FT METAL 832 832
 CC FT METAL 832 832
 CC FT METAL 832 832
 CC SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;
 CC
 CC Query Match 67.3%; Score 35; DB 1; Length 1058;
 CC Best Local Similarity 60.0%; Pred. No. 20;
 CC Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 2 EVVPXGXHY 11
 CC Db 190 EIVPGLNYS 199
 CC
 CC RESULT 2
 CC VLTFF_BPT5 STANDARD; PRT; 1396 AA.
 CC ID VLTFF_BPT5
 CC AC F13390; O48502;
 CC DT 01-JAN-1990 (Rel. 13, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE L-shaped tail fiber protein (LTF protein).
 CC GN LTF
 CC OS Bacteriophage T5.
 CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 CC OC T5-like viruses.
 CC OX NCBI_TaxID=10726;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=95309401; PubMed=7789514;
 CC

RA Kaliman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Kryukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RL FEBS Lett. 366:46-48(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kaliman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kaliman A.V., Kryukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X69460; CAA49220.1; -
 CC DR EMBL; AJ001191; CAA04591.1; -
 CC DR PIR; S01982; S01982.
 KW Late protein.
 FT CONFLICT 986
 FT CONFLICT 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;
 CC
 CC Query Match 67.3%; Score 35; DB 1; Length 1396;
 CC Best Local Similarity 50.0%; Pred. No. 26;
 CC Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 EVVPXGXHY 11
 CC Db 1360 KTIAGDHY 1369
 CC
 CC RESULT 3
 CC YLIK_TYDVA
 CC ID YLIK_TYDVA STANDARD; PRT; 102 AA.
 CC AC Y31619;
 CC DT 01-JUL-1993 (Rel. 26, Created)
 CC DT 01-JUL-1993 (Rel. 26, Last sequence update)
 CC DT 01-OCT-1993 (Rel. 27, Last annotation update)
 CC DE Hypothetical 11.2 kDa protein.
 CC GN VI.
 CC OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
 CC OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 CC OX NCBI_TaxID=31599;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=92188538; PubMed=1546458;
 RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;
 RT "The nucleotide sequence of the infectious cloned DNA component of
 RT tobacco yellow dwarf virus reveals features of geminiviruses
 RT infecting monocotyledonous plants.";
 RL Virology 187:633-642(1992).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81103; AAA47947.1; -
 CC DR PIR; A42452; A42452.

DR InterPro: IPR002621; Gemini_mov.
 DR Pfam: PF01708; Gemini_mov; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 102;
 Best Local Similarity 60.0%; pred. No. 2.7;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXGKHS 11
 :||| |
 Db 7 QVPSGINS 16

RESULT 4

TKNK_BOVIN STANDARD; PRT; 126 AA.

AC P08858;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurokinin B precursor (NKB) (Neuromedin K).
 GN TAC3 OR NKNB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86313713; PubMed=3462746;
 RA Kotani H., Hoshimaru M., Nawa H., Nakanishi S.;
 RT "Structure and gene organization of bovine neuromedin K precursor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; M14351; AAA30723.1;
 DR EMBL; M14347; AAA30723.1; JOINED.
 DR EMBL; M14348; AAA30723.1; JOINED.
 DR EMBL; M14349; AAA30723.1; JOINED.
 DR EMBL; M14350; AAA30723.1; JOINED.
 DR PIR; A25905; A25905.
 DR InterPro: IPR003635; Neurokinin.
 DR InterPro: IPR002040; Tachykinin.
 DR PRODOM: PD020370; Neurokinin; 1.
 DR PROSITE; PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 KW Amidation; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 83
 FT PEPTIDE 86 95 NEUROKININ B.
 FT PROPEP 99 126
 FT MOD_RES 95 95
 SQ SEQUENCE 126 AA; 13871 MW; 446EF433498EC059 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 126;
 Best Local Similarity 66.7%; pred. No. 3.4;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9

Db 28 EQVPGGGH 36

:||| |

RESULT 5

CARB_STAAM STANDARD; PRT; 1057 AA.

AC Q99UR5;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
 GN CARB OR PYRAB OR SAV1203 OR SA1046.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=158878; 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
 RL Lancet 357:1225-1240(2001).
 CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.
 CC -!- COFACTOR: Binds three manganese ions (By similarity).
 CC -!- PATHWAY: Arginine biosynthesis.
 CC -!- PATHWAY: Pyrimidine biosynthesis.
 CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).

 CC -!- SIMILARITY: BELONGS TO THE CARB FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AP003361; BAB57365.1;
 DR EMBL; AP003332; BAB42298.1;
 DR HSSP; P00968; ICS0.
 DR InterPro: IPR005483; Cpsase_L.
 DR InterPro: IPR005479; Cpsase_L_D2.
 DR InterPro: IPR005480; Cpsase_L_D3.
 DR InterPro: IPR005481; Cpsase_L_N.
 DR InterPro: IPR004362; MGS_like.
 DR Pfam; PF02786; Cpsase_L_chain; 2.
 DR Pfam; PF02787; Cpsase_L_D2; 2.
 DR Pfam; PF02787; Cpsase_L_D3; 1.
 DR Pfam; PF02142; MGS; 1.
 DR PRINTS; PR00098; CPSASE.
 DR PROSITE; PS00866; CPSASE_1; 2.
 DR PROSITE; PS00867; CPSASE_2; 2.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 KW ATP-binding; Manganese; Complete proteome.
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 930 1057 ALLOSTERIC DOMAIN.
 FT REPEAT 1 546

```

FT REPEAT      547 1057
FT NP_BIND     153 210
FT NP_BIND     302 352
FT METAL       284 284
FT METAL       298 298
FT METAL       284 284
FT METAL       298 298
FT METAL       300 300
FT METAL       820 820
FT METAL       832 832
SQ SEQUENCE    1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;

Query Match      65.4%; Score 34; DB 1; Length 1057;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHYS 11
   I:| | | |
Db 190 EIVSNGLHYS 199

RESULT 6
CARB_STAAM STANDARD; PRT; 1057 AA.
AC P58940;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
DE phosphate synthetase ammonia chain).
GN CARB OR PYRAB OR MW1086.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamanoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -|- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate
CC -|- COFACTOR: Binds three manganese ions (By similarity).
CC -|- PATHWAY: Arginine biosynthesis.
CC -|- PATHWAY: Pyrimidine biosynthesis; first step.
CC -|- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).
CC -|- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP004825; BAB94951.1; -.
CC DR PROSITE; PS00866; CPSASE.1; 2.
CC DR PROSITE; PS00867; CPSASE.2; 2.
CC
CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
CC ATP-binding; Manganese.
FT DOMAIN      1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN      402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN      547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN      930 1057 ALLOSTERIC DOMAIN.
FT REPEAT      1 546
FT REPEAT      547 1057
FT NP_BIND     153 210 ATP (POTENTIAL).
FT NP_BIND     302 352 ATP (POTENTIAL).

```

```

FT METAL       284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL       298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL       300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL       820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL       832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE    1057 AA; 117185 MW; DBE3B09F9BC6F152 CRC64;

Query Match      65.4%; Score 34; DB 1; Length 1057;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHYS 11
   I:| | | |
Db 190 EIVSNGLHYS 199

RESULT 7
REV_SIVCZ STANDARD; PRT; 124 AA.
AC P17280;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
GN REV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90259077; PubMed=2188136;
RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";
RL Nature 345:356-359(1990).
CC -|- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -|- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -|- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52154; CAA36405.1; -.
CC DR PIR; S09988; VKLJ51.
CC DR HIV; X52154; REVSCPZ.
CC DR InterPro; IPR000625; REV_protein.
CC DR Pfam; PF00424; REV; 1.
CC
CC Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
SQ SEQUENCE    124 AA; 13701 MW; F5877D1BDF65A7B2 CRC64;

Query Match      63.5%; Score 33; DB 1; Length 124;
Best Local Similarity 60.0%; Pred. No. 5.5;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHYS 11
   I:| | | |
Db 107 EIVPAGGNYS 116

RESULT 8
UMEL_YEAST STANDARD; PRT; 460 AA.
AC Q03010; P87330;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Meiosis negative regulator UMEL.

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; Z29372; CRA82563.1; -
CC HSPSP; P00883; IADO.
CC InterPro: IPR000741; Aldolase_I.
CC Pfam; PF00274; glycolytic_enzy; 1.
CC ProDom; PD001128; Aldolase_I; 1.
CC PROSITE; PS00158; ALDOLASE_CLASS_I; 1.
CC Lyase; Schiff base; Glycolysis; Multigene family.
CC INIT_MET 0
CC BINDING 55 55 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
CC BINDING 146 146 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
CC BINDING 229 229 SCHIFF-BASE WITH DIHYDROXYACETONE-P.
CC ACT_SITE 363 363 ESSENTIAL FOR ENHANCED ACTIVITY OF THE
CC ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE
CC AS COMPARED WITH FRUCTOSE 1-PHOSPHATE.
CC SEQUENCE 363 AA; 39500 MW; FC8B4566821E2BD CRC64;

Query Match 61.5%; Score 32; DB 1; Length 363;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXGH 9
||:|:|
DB 189 EVIPDGS 196

RESULT 12

MUTL_BACSU STANDARD; PRT; 627 AA.

ID MUTL_BACSU STANDARD; PRT; 627 AA.

AC P49850;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE DNA mismatch repair protein mutL.

GN MUTL.

OS *Bacillus subtilis*.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

OC NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RC MEDLINE=96349107; PubMed=8760914;

RX Ginetti F., Perego M., Albertini A.M., Galizzi A.;
RT "Bacillus subtilis mutL muts mutL operon: identification, nucleotide
RT sequence and mutagenesis."
RL Microbiology 142:2021-2029(1996).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RC MEDLINE=98044033; PubMed=9384377;

RX Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Deniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Puhlic P., Purnelle B., Rapoport G., Rev M., Reynolds S.,
RA

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN
 CC DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH
 CC REPAIR. MAY ACT AS A "MOLECULAR MATCHMAKER", A PROTEIN THAT
 CC PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE
 CC DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF
 CC BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U7343; AAB19236.1; -;
 DR EMBL: Z99112; CAB13578.1; -;
 DR HSP: P23367; 1BKN.
 DR Subtilist; BG11402; mutL.
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR002099; DNA_mis_repair.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR Pfam: PF01119; DNA_mis_repair; 1.
 DR Pfam: PF02518; HATPase_c; 1.
 DR TIGRams: TIGR00585; mutL; 1.
 DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; 1.
 DR DNA repair; Complete proteome.
 KW SEQUENCE 627 AA; 70431 MW; 068A0509CC265343 CRC64;
 SQ
 Query Match 61.5%; Score 32; DB 1; Length 627;
 Best Local Similarity 54.5%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHY 11
 Db 488 EMIVLTFHYS 498
 RESULT 13
 ETR1_BRAOL STANDARD; PRT; 735 AA.
 ID ETR1_BRAOL
 AC 049230;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ethylene receptor (EC 2.7.3.-).
 ETR1.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3712;
 RN [1]
 RP MEDLINE=98344657; PubMed=9687267;
 RA Chen H.-H., Charny Y.-Y., Yang S.F., Shaw J.-F.;
 RT "Molecular cloning and sequencing of a broccoli cDNA encoding an ETR-
 RT type ethylene receptor.";
 RL (in) Plant Gene Register PGR98-088.
 CC -1- FUNCTION: May act early in the ethylene signal transduction

CC pathway, possibly as an ethylene receptor, or as a regulator of
 CC the pathway (By similarity).
 CC -1- COFACTOR: Binds 1 copper ion per dimer (By similarity).
 CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum (By similarity).
 CC -1- PTM: Activation probably requires a transfer of a phosphate group
 CC between a His in the transmitter domain and a Asp of the receiver
 CC domain (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE ETHYLENE RECEPTOR FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF047476; AAC39497.1; -;
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR003661; His_kinA.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR InterPro: IPR001789; Response_reg.
 DR Pfam: PF01590; GAF; 1.
 DR Pfam: PF02518; HATPase_c; 1.
 DR Pfam: PF00072; response_reg; 1.
 DR Pfam: PF00512; signal; 1.
 DR ProDom: PD000039; Response_reg; 1.
 DR SMART: SM00065; GAF; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00388; HisKA; 1.
 DR SMART: SM00448; REC; 1.
 DR PROSITE: PS0109; HIS_KIN; 1.
 DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
 DR Receptor; Sensory transduction; Phosphorylation; Transmembrane;
 KW Transferase; Kinase; Copper; Metal-binding.
 FT DOMAIN 350 586 HISTIDINE KINASE.
 FT DOMAIN 609 726 RESPONSE REGULATORY.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 54 74 POTENTIAL.
 FT TRANSMEM 92 112 POTENTIAL.
 FT METAL 65 65 COPPER (BY SIMILARITY).
 FT METAL 69 69 COPPER (BY SIMILARITY).
 FT DISULFID 4 4 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 6 6 INTERCHAIN (BY SIMILARITY).
 FT MOD_RES 353 353 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 657 657 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 735 AA; 82240 MW; 520B77291CF2BCC6 CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 735;
 Best Local Similarity 62.5%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGXHY 10
 Db 499 VVPTGSHF 506
 RESULT 14
 ETR1_ARATH STANDARD; PRT; 738 AA.
 ID ETR1_ARATH
 AC P49333;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ethylene receptor (EC 2.7.3.-).
 GN ETR1 OR AT1G66340 OR T27F4.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurons II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTANTS ETR1-1; ETR1-2; ETR1-3 AND ETR1-4.
 RX MEDLINE=94024007; PubMed=8211181;
 RA Chang C., Kwok S.F., Blecker A.B., Meyerowitz E.M.;
 RA "Arabidopsis ethylene-response gene ETR1: similarity of product to
 RT two-component regulators.";
 RL Science 262:539-544(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cy Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan O., Chen Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Millischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP DISULFIDE BONDS, AND MUTAGENESIS OF CYS-4; CYS-6 AND CYS-99.
 RX PubMed=7759498;
 RA Schaller G.E., Ladd A.N., Lanahan M.B., Spanbauer J.M., Blecker A.B.;
 RT "The ethylene response mediator ETR1 from Arabidopsis forms a
 RT disulfide-linked dimer.";
 RL J. Biol. Chem. 270:12526-12530(1995).
 RN [4]
 RP MUTAGENESIS OF CYS-4; CYS-6; CYS-65 AND CYS-99.
 RX PubMed=8525372;
 RA Schaller G.E., Blecker A.B.;
 RT "Ethylene-binding sites generated in yeast expressing the Arabidopsis
 RT ETR1 gene.";
 RL Science 270:1809-1811(1995).
 RN [5]
 RP COFACTOR, AND MUTAGENESIS OF GLU-38; CYS-65; HIS-69; MET-87;
 RP MET-104 AND HIS-107.
 RX PubMed=9974395;
 RA Rodriguez F.I., Esch J.J., Hall A.E., Binder B.M., Schaller G.E.,
 RA Blecker A.B.;
 RT "A copper cofactor for the ethylene receptor ETR1 from Arabidopsis.";
 RL Science 283:996-998(1999).
 RN [6]
 RP SUBCELLULAR LOCATION.
 RX PubMed=11916973;
 RA Chen Y.-F., Randlett M.D., Findell J.L., Schaller G.E.;
 RT "Localization of the ethylene receptor ETR1 to the endoplasmic
 RT reticulum of Arabidopsis.";
 RL J. Biol. Chem. 277:15961-15966(2002).
 CC [1]- FUNCTION: May act early in the ethylene signal transduction
 CC pathway, possibly as an ethylene receptor, or as a regulator of
 CC the pathway.
 CC [1]- COFACTOR: Binds 1 copper ion per dimer.
 CC [1]- SUBUNIT: Homodimer; disulfide-linked.
 CC [1]- SUBCELLULAR LOCATION: Endoplasmic reticulum.
 CC [1]- TISSUE SPECIFICITY: Leaves, roots, stems, flowers and seedlings.
 CC [1]- PTM: Activation probably requires a transfer of a phosphate group
 CC between a His in the transmitter domain and a Asp of the receiver
 CC domain (By similarity).
 CC [1]- SIMILARITY: BELONGS TO THE ETHYLENE RECEPTOR FAMILY.
 CC [1]- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.

CC [1]- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L24119; AAA70047.1; -;
 CC EMBL; AC020665; AAG52169.1; -;
 CC InterPro; IPR003594; ATPbind_ATPase.
 CC InterPro; IPR003018; GAF.
 CC InterPro; IPR004359; HIS_KIN_sig.
 CC InterPro; IPR003661; His_kinA.
 CC InterPro; IPR001789; Response_reg.
 CC Pfam; PF00072; response_reg; 1.
 CC Pfam; PF00512; signal; 1.
 CC Pfam; PF01590; GAF; 1.
 CC Pfam; PF02518; HATPase_c; 1.
 CC ProDom; PD000039; Response_reg; 1.
 CC SMART; SM00065; GAF; 1.
 CC SMART; SM00387; HATPase_c; 1.
 CC SMART; SM00388; HSKA; 1.
 CC SMART; SM00448; REC; 1.
 CC PROSITE; PS50109; HIS_KIN; 1.
 CC PROSITE; PS50110; RESPONSE_REGULATORY; 1.
 CC Receptor; Sensory transduction; Phosphorylation; Transmembrane;
 CC Transferrase; Kinase; Copper; Metal-binding.
 KW TRANSFERASE; KINASE; COPPER; METAL-BINDING.
 FT DOMAIN 350 585
 FT DOMAIN 611 729
 FT TRANSMEM 23 43
 FT TRANSMEM 53 73
 FT TRANSMEM 92 112
 FT METAL 65 65
 FT METAL 69 69
 FT DISULFID 4 4
 FT DISULFID 6 6
 FT MOD_RES 353 353
 FT MOD_RES 659 659
 FT MUTAGEN 4 4
 FT MUTAGEN 6 6
 FT MUTAGEN 31 31
 FT MUTAGEN 38 38
 FT MUTAGEN 62 62
 FT MUTAGEN 65 65
 FT MUTAGEN 69 69
 FT MUTAGEN 79 79
 FT MUTAGEN 99 99
 FT MUTAGEN 102 102
 FT MUTAGEN 104 104
 FT MUTAGEN 107 107
 SQ SEQUENCE 738 AA; 82565 MW; D6ED3C4BBA87E96E CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 738;
 Best Local Similarity 62.5%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGXHY 10
 Db 498 VVPTGSHF 505
 RESULT 15
 CENC_MOUSE STANDARD; PRT; 906 AA.
 ID CENC_MOUSE
 AC P49452;

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Centromere protein C (CENP-C) (Centromere autoantigen C).
 GN CENPC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss;
 RX MEDLINE=95048391; PubMed=7959789;
 RA McKay S., Thomson E., Cooke H.;
 RT "Sequence homologies and linkage group conservation of the human and
 mouse CenpC genes.";
 RL Genomics 22:36-40(1994).
 CC -!- FUNCTION: COMPONENT OF THE INNER KINETOCHORE PLATE. REQUIRED FOR
 CC NORMAL KINETOCHORE ASSEMBLY.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U03113; AAC04314.1; -.
 DR MGD; MGI:99700; Cenpc.
 KW Chromosomal protein; Nuclear protein; DNA-binding; Phosphorylation;
 KW Centromere.
 SQ SEQUENCE 906 AA; 102225 MW; 43D529B202E9D71E CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 906;
 Best Local Similarity 62.5%; Pred. No. 72;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 4 VPXGXHYIS 11
 ||| |||
 Db 881 VPSGNHYN 888
 RESULT 16
 YQS6_CAEEL STANDARD; PRT; 1374 AA.
 ID YQS6_CAEEL
 AC Q09541;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical subtilase-type proteinase F21H12.6 in chromosome II
 DE (EC 3.4.21.-).
 GN F21H12.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Favell T.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U23176; AAC46718.1; -.
 DR HSP; P00782; ISUP.
 DR MEROPS; S08 090; -.
 DR WormPep; F21H12.6; CE01917.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hypothetical protein; Hydrolase; Serine protease.
 FT ACT_SITE 261 261 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 314 314 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 499 499 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 1374 AA; 151088 MW; 6937C862AED1E7FB CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 1374;
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXHYIS 11
 ||| ||| |||
 Db 306 EIVVPSGAHGS 316
 RESULT 17
 RON_MOUSE STANDARD; PRT; 1378 AA.
 ID RON_MOUSE
 AC Q62190; Q62555;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Macrophage-stimulating protein receptor precursor (EC 2.7.1.112)
 DE (MSP receptor) (p185-Ron) (Stem cell-derived tyrosine kinase).
 GN MSTIR OR RON OR STK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94250897; PubMed=8193352;
 RA Iwama A., Okano A., Sudo T., Matsuda Y., Suda T.;
 RT "Molecular cloning of a novel receptor tyrosine kinase gene, STK,
 RT derived from enriched hematopoietic stem cells.";
 RL Blood 83:3160-3169(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=98127434; PubMed=9467940;
 RA Waltz S.E., Toms C.L.V., McDowell S.A., Clay L.A., Muraoka R.S.,
 RA Air E.L., Sun W.Y., Thomas M.B., Degen S.J.F.;
 RT "Characterization of the mouse Ron/Stk receptor tyrosine kinase
 RT gene.";
 RL Oncogene 16:27-42(1998).
 CC -!- FUNCTION: RECEPTOR FOR MACROPHAGE STIMULATING PROTEIN (MSP). HAS A
 CC TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN AND A BETA
 CC CHAIN WHICH ARE DISULFIDE LINKED (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- PTM: THE TWO SUBUNITS ARE PROTEOLYTICALLY DERIVED (BY SIMILARITY).
 CC -!- PTM: PHOSPHORYLATED IN RESPONSE TO LIGAND BINDING (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

or send an email to license@isb-sib.ch).

```

CC CC EMBL; X74736; CAA52754.1; -.
DR DR EMBL; U65949; AAC39953.1; -.
DR DR HSSP; P11362; 1FGK.
DR DR MGI; 99614; Mst1r.
DR DR InterPro; IPR000719; Euk_pkinase.
DR DR InterPro; IPR002909; IPT_TIG.
DR DR InterPro; IPR002165; Plexin_repeat.
DR DR InterPro; IPR001627; Sema.
DR DR InterPro; IPR002290; Ser_thr_pkinase.
DR DR InterPro; IPR001245; Tyr_pkinase.
DR DR Pfam; PF00069; pkinase; 1.
DR DR Pfam; PF01403; Sema; 1.
DR DR Pfam; PF01437; PSI; 1.
DR DR Pfam; PF01833; TIG; 3.
DR DR ProDom; PD000001; Euk_pkinase; 1.
DR DR SMART; SM00220; S_TKC; 1.
DR DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR DR Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; Signal; Phosphorylation; ATP-binding.
FT FT SIGNAL 1 23
FT FT CHAIN 24 1378
FT FT CHAIN 25 305
FT FT CHAIN 311 1378
FT FT DOMAIN 25 960
FT FT TRANSMEM 961 981
FT FT DOMAIN 982 1378
FT FT DOMAIN 60 509
FT FT DOMAIN 1059 1322
FT FT NP_BIND 1065 1073
FT FT BINDING 1091 1091
FT FT ACT_SITE 1185 1185
FT FT MOD_RES 1330 1330
FT FT MOD_RES 1337 1337
FT FT CARBOHYD 91 91
FT FT CARBOHYD 391 391
FT FT CARBOHYD 460 460
FT FT CARBOHYD 490 490
FT FT CARBOHYD 656 656
FT FT CARBOHYD 722 722
FT FT CARBOHYD 845 845
FT FT CARBOHYD 901 901
FT FT CONFLICT 713 714
FT FT CONFLICT 719 719
SQ SEQUENCE 1378 AA; 150538 MW; FC5F932377B57009 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 1378;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9
DB 601 DDVVPGEH 609

RESULT 18
SZ05_RAT STANDARD; PRT; 130 AA.
AC P97855;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine B5 precursor (CXCL5) (Cytokine LIX).
GN SCYB5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=101116;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley;
RA Kelner G.S., Maciejewski-Lenoir D., Lee E.D., Maki R.A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY PARTICIPATE IN THE RECRUITMENT OF INFLAMMATORY CELLS
CC BY INJURED OR INFECTED TISSUE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXC).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).


EMBL; U90448; AAB61460.1; -.
HSSP; P10889; 1MI2.
DR InterPro; IPR001089; CXC_chemokine_smll.
DR InterPro; IPR001811; Chemokine_IL8.
DR Pfam; PF00048; IL8; 1.
DR PRINTS; PR00437; SMALLCYTKCXC.
DR SMART; SM00199; SCY; 1.
DR PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Signal.
FT FT SIGNAL 1 37
FT FT CHAIN 38 130
FT FT DISULFID 50 76
FT FT DISULFID 52 93
SQ SEQUENCE 130 AA; 14263 MW; C00F6B3605524F4E CRC64;

Query Match 59.6%; Score 31; DB 1; Length 130;
Best Local Similarity 62.5%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXH 9
DB 68 EVIPAGPH 75

RESULT 19
MSRA_VIBCH STANDARD; PRT; 212 AA.
ID MSRA_VIBCH AC Q9KP30;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide methionine sulfoxide reductase msra (EC 1.8.4.6) (Protein-
DE methionine-S-oxide reductase) (Peptide Met(O) reductase).
GN MSRA OR VC2549.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=686;
RN RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Has an important function as a repair enzyme for
CC proteins that have been inactivated by oxidation. Catalyzes the
CC reversible oxidation-reduction of methionine sulfoxide in proteins
CC to methionine (By similarity).


```

CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC protein L-methionine S-oxide + reduced thioredoxin.
CC -1- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004324; AAF95690.1; -
CC TIGR; VC2549; -
CC InterPro; IPR002569; PMSR.
CC Pfam; PF01625; PMSR; 1.
CC ProDom; PD003489; PMSR; 1.
CC TIGRFAMs; TIGR00401; msrA; 1.
CC OXidoreductase; Complete proteome.
CC ACT SITE 51 51 BY SIMILARITY.
CC SEQUENCE 212 AA; 23360 MW; AD50686E8BFF13016 CRC64;
CC
CC Query Match 59.6%; Score 31; DB 1; Length 212;
CC Best Local Similarity 44.4%; Pred. No. 26;
CC Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 2 EVVPGXHY 10
CC |::| |:
CC 171 EILPAGAYY 179
CC
CC RESULT 20
CC YHAI_CRYPA STANDARD; PRT; 319 AA.
CC ID YHAI_CRYPA
CC AC P10941;
CC DT 01-JUL-1989 (Rel. 11, Created)
CC DT 01-JUL-1989 (Rel. 11, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein 1 in hypovirulence-associated DS-RNA genetic
CC element [contains: P29 proteinase].
CC OS Cryphonectria parasitica (Chestnut blight fungus) (Endothia
CC parasitica).
CC OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC OC Diaporthales; Valsaceae; Cryphonectria.
CC OX NCBI_TaxID=5116;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RC STRAIN=EP713;
CC RX MEDLINE=89251594; PubMed=2721496;
CC RA Rae B.P., Hillman B.I., Tartaglia J., Nuss D.L.;
CC RT "Characterization of double-stranded RNA genetic elements associated
CC with biological control of chestnut blight: organization of terminal
CC RT domains and identification of gene products.";
CC RL EMO J. 8:657-663(1989).
CC -1- MISCELLANEOUS: DOUBLE-STRANDED RNA GENETIC ELEMENTS ARE ASSOCIATED
CC WITH BIOLOGICAL CONTROL OF THE FUNGAL DISEASE CHESTNUT BLIGHT.
CC THIS DS-RNA ARE ASSOCIATED WITH HYPOVIRULENCE. THEY ARE LOCALIZED
CC IN THE CYTOPLASM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C7.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14524; CAA32666.1; -
CC PIR; S03833; S03833.
CC MEROPS; C07.001; -
CC InterPro; IPR002704; Peptidase_C7.

DR Pfam; PF01830; Peptidase_C7; 1.
DR ProDom; PD040949; Peptidase_C7; 1.
KW Hypothetical protein; Hydrolase; Thiol protease. CRC64;
SQ SEQUENCE 319 AA; 35443 MW; A1F5F775F9ACE7A CRC64;
CC
CC Query Match 59.6%; Score 31; DB 1; Length 319;
CC Best Local Similarity 85.7%; Pred. No. 40;
CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 EEVVPXG 7
CC |::| |:
CC 31 EEVVPAG 37
CC
CC RESULT 21
CC ALPB_HUMAN STANDARD; PRT; 363 AA.
CC ID ALPB_HUMAN
CC AC P05062; Q13741; Q13742;
CC DT 13-AUG-1987 (Rel. 05, Created)
CC DT 13-AUG-1987 (Rel. 05, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Fructose-bisphosphate aldolase B (EC 4.1.2.13) (Liver-type aldolase).
CC GN ALDOB OR ALDB.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=85037920; PubMed=6548561;
CC RA Paolella G., Santamaria R., Izso P., Costanzo P., Salvatore F.;
CC RT "Isolation and nucleotide sequence of a full-length cDNA coding for
CC RT aldolase B from human liver.";
CC RL Nucleic Acids Res. 12:7401-7410(1984).
CC [2]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=85269627; PubMed=2410860;
CC RA Sakakibara M., Mukai T., Yatsuki H., Hori K.;
CC RT "Human aldolase isozyme gene: the structure of multispecies aldolase
CC RT B mRNAs.";
CC RL Nucleic Acids Res. 13:5055-5069(1985).
CC [3]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=84194081; PubMed=6585824;
CC RA Rottmann W.H., Toland D.R., Penhoet E.E.;
CC RT "Complete amino acid sequence for human aldolase B derived from cDNA
CC RT and genomic clones.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 81:2738-2742(1984).
CC [4]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=88139255; PubMed=2830249;
CC RA Mukai T., Yatsuki H., Arai Y., Joh K., Matsuhashi S., Hori K.;
CC RT "Human aldolase B gene: characterization of the genomic aldolase B
CC RT gene and analysis of sequences required for multiple
CC RT polyadenylations.";
CC RL J. Biochem. 102:1043-1051(1987).
CC [5]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=86284196; PubMed=3016456;
CC RA Toland D.R., Penhoet E.E.;
CC RT "Characterization of the human aldolase B gene.";
CC RL Mol. Biol. Med. 3:245-264(1986).
CC [6]
CC RN SEQUENCE OF 238-364 FROM N.A.
CC RX MEDLINE=84104270; PubMed=6689266;
CC RA Besmond C., Dreyfus J.-C., Gregori C., Frain M., Zakin M.M.,
CC RA Sala Trepal J., Kahn A.;
CC RT "Nucleotide sequence of a cDNA clone for human aldolase B.";
CC RL Biochem. Biophys. Res. Commun. 117:601-609(1983).
CC [7]
CC RN REVIEW ON VARIANTS.
CC RX MEDLINE=96121610; PubMed=8535439;
CC RA Toland D.R.;

RT "Molecular basis of hereditary fructose intolerance: mutations and polymorphisms in the human aldolase B gene.";

RL Hum. Mutat. 6:210-218(1995).

RN [8]

RP VARIANT HFI.

RX MEDLINE=88253408; PubMed=3383242;

RA Cross N.C.P., Tolan D.R., Cox T.M.;

RT "Catalytic deficiency of human aldolase B in hereditary fructose intolerance caused by a common missense mutation.";

RL Cell 53:881-885(1988).

RN [9]

RP VARIANT HFI.

RX MEDLINE=90135738; PubMed=1967768;

RA Cross N.C.P., de Franchis R., Sebastio G., Dazzo C., Tolan D.R.,

RA Gregori C., Odievre M., Vidailhet M., Romano V., Mascali G.,

RA Romano C., Musumeci S., Steinmann B., Gitzelmann R., Cox T.;

RT "Molecular analysis of aldolase B genes in hereditary fructose intolerance.";

RL Lancet 335:306-309(1990).

RN [10]

RP VARIANT HFI ARG-134.

RX MEDLINE=94131206; PubMed=8299883;

RA Brooks C.C., Tolan D.R.;

RT "A partially active mutant aldolase B from a patient with hereditary fructose intolerance.";

RL FASEB J. 8:107-113(1994).

RN [11]

RP VARIANT HFI ARG-147.

RX MEDLINE=95233426; PubMed=7717389;

RA Ali M., Cox T.M.;

RT "Diverse mutations in the aldolase B gene that underlie the prevalence of hereditary fructose intolerance.";

RL Am. J. Hum. Genet. 56:1002-1005(1995).

RN [12]

RP VARIANT HFI PRO-256.

RX MEDLINE=94214447; PubMed=8162030;

RA Ali M., Sebastio G., Cox T.M.;

RT "Identification of a novel mutation (Leu 256-->Pro) in the human aldolase B gene associated with hereditary fructose intolerance.";

RL Hum. Mol. Genet. 3:203-204(1994).

RN [13]

RP VARIANT HFI LYS-334.

RX MEDLINE=90245619; PubMed=2336380;

RA Cross N.C.P., Stojanov L.M., Cox T.M.;

RT "A new aldolase B variant, N334K, is a common cause of hereditary fructose intolerance in Yugoslavia.";

RL Nucleic Acids Res. 18:1925-1925(1990).

CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glyceralone phosphate + D-glyceraldehyde 3-phosphate.

CC -1- PATHWAY: glycolysis; sixth step.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- DISEASE: DEFECTS IN ALDOB ARE THE CAUSE OF HEREDITARY FRUCTOSE INTOLERANCE (HFI); AN AUTOSOMAL RECESSIVE DISEASE THAT RESULTS IN AN INABILITY TO METABOLIZE FRUCTOSE AND RELATED SUGARS. COMPLETE EXCLUSION OF FRUCTOSE RESULTS IN DRAMATIC RECOVERY; HOWEVER, IF NOT TREATED PROPERLY, HFI SUBJECTS SUFFER EPISODES OF HYPOGLYCEMIA, GENERAL ILL HEALTH, AND RISK OF DEATH THE REMAINDER OF LIFE.

CC -1- MISCELLANEOUS: IN VERTEBRATES, THREE FORMS OF THIS UBIQUITOUS GLYCOLYTIC ENZYME ARE FOUND, ALDOLASE A IN MUSCLE, ALDOLASE B IN LIVER, & ALDOLASE C IN BRAIN.

CC -1- SIMILARITY: BELONGS TO CLASS I FRUCTOSE-BISPHOSPHATE ALDOLASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X02747; CAA26526.1; ;

DR EMBL; X01098; CAA25572.1; ;

DR EMBL; K01177; AAB59377.1; ;

DR EMBL; D00183; BAA00125.1; ;

DR EMBL; D00176; BAA00125.1; JOINED.

DR EMBL; D00177; BAA00125.1; JOINED.

DR EMBL; D00178; BAA00125.1; JOINED.

DR EMBL; D00179; BAA00125.1; JOINED.

DR EMBL; D00180; BAA00125.1; JOINED.

DR EMBL; D00181; BAA00125.1; JOINED.

DR EMBL; D00182; BAA00125.1; JOINED.

DR EMBL; M15656; AAA51691.1; ;

DR EMBL; X00270; CAA25072.1; ;

DR EMBL; X00270; CAA25072.1; JOINED.

DR PIR; A23788; ADHUB.

DR HSP; P00883; LADOB.

DR Genew; HGNC:417; ALDOB.

DR MIM; 229600; ;

DR InterPro; IPR000741; Aldolase_I.

DR Pfam; PF00274; glycolytic_enzy; 2.

DR ProDom; PD001128; Aldolase_I; 1.

DR PROSITE; PS00158; ALDOLASE_CLASS_I; 1.

DR Lyase; Schiff base; Glycolysis; Multigene family; Disease mutation.

FT INIT_MET 0

FT BINDING 55

FT BINDING 146

FT BINDING 229

FT ACT_SITE 363

FT 134

FT VARIANT 134

FT 147

FT VARIANT 147

FT 149

FT VARIANT 149

FT 174

FT VARIANT 174

FT 256

FT VARIANT 256

FT 303

FT VARIANT 303

FT 334

FT VARIANT 334

FT 337

FT VARIANT 337

FT 53

FT CONFLICT 53

FT 249

FT CONFLICT 249

FT 277

FT CONFLICT 277

FT 308

FT CONFLICT 308

FT 347

FT CONFLICT 347

FT SEQUENCE 363 AA; 39342 MW; B3371B1012B2DBA8 CRC64;

SQ

Query Match 59.6%; Score 31; DB 1; Length 363;

Best Local Similarity 62.5%; Pred. No. 45;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXGH 9

Db 189 EVIIPGDH 196

RESULT 22

ALFB_RABIT

ID ALFB_RABIT STANDARD; PRT; 363 AA.

AC P79226;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Fructose-bisphosphate aldolase B (EC 4.1.2.13) (Liver-type aldolase).

GN ALDOB OR ALDB.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN NCBI_TaxID=9986;
 [1]
 RP SEQUENCE FROM N.A.
 RA Berardini T.Z., Amsden A.B., Penhoet E.E., Tolan D.R.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
 CC phosphate + D-glyceraldehyde 3-phosphate.
 CC -1- PATHWAY: Glycolysis; sixth step.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- MISCELLANEOUS: IN VERTEBRATES, THREE FORMS OF THIS UBIQUITOUS
 CC GLYCOLYTIC ENZYME ARE FOUND, ALDOLASE A IN MUSCLE, ALDOLASE B IN
 CC LIVER, & ALDOLASE C IN BRAIN.
 CC -1- SIMILARITY: BELONGS TO CLASS I FRUCTOSE-BISPHOSPHATE ALDOLASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U85645; AAB42087.1; -
 DR HSP; P00883; IADO.
 DR InterPro: IPR000741; Aldolase_I.
 DR Pfam; PF00274; glycolytic_enz; 1.
 DR ProDom; PD001128; Aldolase_I; 1.
 DR PROSITE; PS00158; ALDOLASE_CLASS_I; 1.
 KW Lyase; Schiff base; Glycolysis; Multigene family.
 FT INIT_MET 0 BY SIMILARITY.
 FT BINDING 55 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 FT BINDING 146 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 FT BINDING 229 SCHIFF-BASE WITH DIHYDROXYACETONE-P.
 FT ACT_SITE 363 ESSENTIAL FOR ENHANCED ACTIVITY OF THE
 FT ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE.
 FT AS COMPARED WITH FRUCTOSE 1-PHOSPHATE.
 SQ SEQUENCE 363 AA; 39474 MW; 0A7185A7E89F436 CRC64;
 Query Match 59.6%; Score 31; DB 1; Length 363;
 Best Local Similarity 62.5%; Pred. NO. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EVVPXGXH 9
 Db 189 EVIPDGDH 196
 ||:| | |
 RESULT 23
 THIK_ECOLI STANDARD; PRT; 387 AA.
 ID THIK_ECOLI
 AC P21151; P78130;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-ketoacyl-CoA thiolase (EC 2.3.1.16) (Fatty oxidation complex beta
 DE subunit) (beta-ketothiolase) (Acetyl-CoA acyltransferase).
 GN FADA OR OLDA OR B3845.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-92358234; PubMed-1379743;
 RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
 RT "Analysis of the Escherichia coli genome: DNA sequence of the region
 RL from 84.5 to 86.5 minutes."
 RN Science 257:771-778(1992).
 [2]

RP SEQUENCE FROM N.A.
 RA MEDLINE-91035260; PubMed-1699931;
 RX Dirusso C.C.;
 RT "Primary sequence of the Escherichia coli fadBA operon, encoding the
 RT fatty acid-oxidizing multienzyme complex, indicates a high degree of
 RT homology to eucaryotic enzymes."
 RL J. Bacteriol. 172:6459-6468(1990).
 [3]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
 RC STRAIN-K12;
 RX MEDLINE-90285166; PubMed-2191949;
 RA Yang S.-Y., Yang X.-Y.H., Healy-Louie G., Schulz H., Elzinga M.;
 RT "Nucleotide sequence of the fadA gene. Primary structure of
 RT 3-ketoacyl-coenzyme A thiolase from Escherichia coli and the
 RT structural organization of the fadAB operon."
 RL J. Biol. Chem. 265:10424-10429(1990).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RX MEDLINE-90370500; PubMed-2204034;
 RA Nakahigashi K., Inokuchi H.;
 RT "Nucleotide sequence of the fadA and fadB genes from Escherichia
 RT coli."
 RL Nucleic Acids Res. 18:4937-4937(1990).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1232-1244(1997).
 CC -1- FUNCTION: FADB AND FADA ARE THE ALPHA AND BETA SUBUNITS OF THE
 CC MULTIFUNCTIONAL ENZYME COMPLEX OF THE FATTY ACID DEGRADATION
 CC CYCLE.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + acetyl-CoA = CoA + 3-oxoacyl-CoA.
 CC -1- PATHWAY: Fatty acid beta-oxidation cycle.
 CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE THIOLEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M87049; AAA67642.1; -
 DR EMBL; M59368; AAA23751.1; -
 DR EMBL; X52837; CAB40810.1; -
 DR EMBL; M74164; AAA62778.1; -
 DR EMBL; AE000460; AAC76848.1; -
 DR PIR; A35436; XUCF.
 DR PIR; JQ0655; JQ0655.
 DR PIR; JY0109; JY0109.
 DR PIR; S30736; S30736.
 DR HSP; P27796; IAFY.
 DR ECO2DBASE; H038.6; 6TH EDITION.
 DR ECO2DBASE; H038.7; 6TH EDITION.
 DR EcoGene; EG10278; fada.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam; PF00108; thiolase; 1.
 DR PROSITE; PS00098; THIOLEASE_C; 1.
 DR PROSITE; PS00099; THIOLEASE_1; 1.
 DR PROSITE; PS00099; THIOLEASE_3; 1.
 DR PROSITE; PS00737; THIOLEASE_2; 1.
 KW Fatty acid metabolism; Transferase; Acyltransferase;
 KW Complete proteome.
 ACT_SITE 91 SUBSTRATE BINDING (BY SIMILARITY).
 FT ACT_SITE 91

FT ACT SITE 373 373 BASE (BY SIMILARITY).
 FT CONFLICT 37 37 T -> S (IN REF. 2 AND 3).
 FT CONFLICT 119 119 E -> G (IN REF. 4).
 FT CONFLICT 371 374 TMC1 -> DGCVS (IN REF. 3).
 SQ SEQUENCE 387 AA; 40890 MW; CB0FF8EF4597D8BA CRC64;

Query Match 59.6%; Score 31; DB 1; Length 387;
 Best Local Similarity 50.0%; Pred. No. 48;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXH 9

Db 188 EIIPTGGH 195

RESULT 24

SLS1_YARLI
 ID SLS1_YARLI STANDARD; PRT; 426 AA.
 AC Q99158;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SLS1 protein precursor.
 GN SLS1.
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;

SEQUENCE FROM N.A.

STRAIN-ATCC 20460 / W29;

MEDLINE=96216076; PubMed=9662639;

Boisrame A., Beckerich J.-M., Gaillardin C.;

"Slisip, an endoplasmic reticulum component, is involved in the

protein translocation process in the yeast Yarrowia lipolytica.";

J. Biol. Chem. 271:11668-11675(1996).

-I- FUNCTION: INVOLVED IN THE PROTEIN TRANSLOCATION PROCESS. MAY

INTERACT DIRECTLY WITH TRANSLATING POLYPEPTIDES TO FACILITATE

THEIR TRANSFER AND/OR HELP THEIR FOLDING IN THE ER. IT IS NOT

REQUIRED FOR VIABILITY BUT IS ESSENTIAL FOR OPTIMAL GROWTH AT

ELEVATED TEMPERATURES.

-I- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; Z50154; CAA90516.1; .

InterPro; IPR000886; ER_target.

PROSITE; PS00014; ER_TARGET; 1.

Endoplasmic reticulum; Signal.

SIGNAL 1 17 POTENTIAL.

CHAIN 18 426 SLS1 PROTEIN.

FT SITE 423 426 PREVENT SECRETION FROM ER (POTENTIAL).

SQ SEQUENCE 426 AA; 47201 MW; 0ACD7EF17540B8E2 CRC64;

Query Match

Best Local Similarity 59.6%; Score 31; DB 1; Length 426;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9

Db 52 DQVIPAGLH 60

RESULT 25

TOLB_HAEIN STANDARD; PRT; 427 AA.

ID TOLB_HAEIN

AC P44677; P94811;

DT

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE TOLB protein precursor.

GN TOLB OR HI0382.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]

SEQUENCE FROM N.A.

STRAIN-Rd / KW20 / ATCC 51907;

MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,

Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

Utterback T.R., Hanna M.C., Nguyen D.T., Saudake D.M., Brandon R.C.,

Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

Venter J.C.;

"Whole-genome random sequencing and assembly of Haemophilus influenzae

Rd.";

Science 269:496-512(1995).

[2]

SEQUENCE FROM N.A.

STRAIN=1479;

RC MEDLINE=97080550; PubMed=8921895;

RA Sen K., Sikkema D.J., Murphy T.F.;

"Isolation and characterization of the Haemophilus influenzae tolQ,

tolR, tolA and tolB genes.";

Gene 178:75-81(1996).

CC -I- FUNCTION: INVOLVED IN THE TONB-INDEPENDENT UPTAKE OF PROTEINS

(BY SIMILARITY).

-I- SUBCELLULAR LOCATION: Periplasmic (Potential).

-I- SIMILARITY: BELONGS TO THE TOLB FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U32722; AAC22040.1; .

EMBL; U32470; AAC44597.1; .

HSSP; P19935; ICRZ.

TIGR; HI0382; .

KW Transport; Protein transport; Periplasmic; Signal; Complete proteome.

SIGNAL 1 23 POTENTIAL.

CHAIN 24 427 TOLB PROTEIN.

FT VARIANT 6 6 R -> H (IN STRAIN 1479).

FT VARIANT 14 14 V -> I (IN STRAIN 1479).

FT VARIANT 17 19 VGS -> ITH (IN STRAIN 1479).

FT VARIANT 21 21 A -> V (IN STRAIN 1479).

FT VARIANT 29 79 R -> H (IN STRAIN 1479).

FT VARIANT 71 79 T -> A (IN STRAIN 1479).

FT VARIANT 129 129 A -> G (IN STRAIN 1479).

FT VARIANT 160 160 A -> T (IN STRAIN 1479).

FT VARIANT 237 237 A -> N (IN STRAIN 1479).

FT VARIANT 322 322 S -> N (IN STRAIN 1479).

FT VARIANT 326 326 A -> V (IN STRAIN 1479).

FT VARIANT 328 328 A -> S (IN STRAIN 1479).

SQ SEQUENCE 427 AA; 44967 MW; 0882201AE9254B9 CRC64;

Query Match 59.6%; Score 31; DB 1; Length 427;

Best Local Similarity 60.0%; Pred. No. 54;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHS 11

Db 103 QVPSGNGYS 112

Search completed: June 10, 2003, 13:40:24
Job time : 5.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-44
Perfect score: 52
Sequence: 1 EEVVPXGXHYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	69.2	208	2	Q46486
2	35	67.3	233	10	Q40479
3	35	67.3	237	10	Q9LW50
4	35	67.3	285	16	Q98HU6
5	35	67.3	308	16	Q9X2E2
6	35	67.3	317	9	Q38317
7	35	67.3	350	17	Q27679
8	35	67.3	678	12	Q9EIX6
9	35	67.3	1028	16	Q8YJ11
10	35	67.3	1063	16	Q8RG86
11	34	65.4	279	16	Q9RXN9
12	34	65.4	322	17	Q9HLH8
13	34	65.4	495	5	Q16912
14	34	65.4	534	17	Q29966
15	34	65.4	822	2	Q9ZHG7
16	34	65.4	825	16	Q99XV4

17	34	65.4	1057	16	Q99UR5
18	34	65.4	1442	17	Q96YH5
19	33	63.5	78	6	Q9XST4
20	33	63.5	139	2	Q57489
21	33	63.5	156	16	Q9PC35
22	33	63.5	217	4	Q00404
23	33	63.5	252	16	Q8YWP1
24	33	63.5	299	4	Q9UEE9
25	33	63.5	412	5	Q08523
26	33	63.5	480	5	Q95NK1
27	33	63.5	510	10	Q9SA71
28	33	63.5	722	12	Q9QTC7
29	33	63.5	748	12	Q9YR01
30	33	63.5	795	5	Q9V8K6
31	33	63.5	802	16	Q97QM8
32	33	63.5	816	2	Q9AHT9
33	33	63.5	819	2	Q9ANY3
34	33	63.5	819	10	Q9AVK4
35	33	63.5	819	16	Q97QM9
36	33	63.5	839	16	Q9ANY2
37	33	63.5	844	2	Q9AG74
38	33	63.5	848	16	Q8P6S4
39	33	63.5	1039	16	Q9ANY1
40	32	61.5	132	2	Q9SIN2
41	32	61.5	135	4	Q9UBD2
42	32	61.5	143	17	Q8TX62
43	32	61.5	162	16	Q53667
44	32	61.5	201	16	Q9L2F0
45	32	61.5	225	10	Q40129
46	32	61.5	246	5	Q9N366
47	32	61.5	280	5	Q9VMN9
48	32	61.5	286	10	Q9ZNT7
49	32	61.5	327	16	Q8UJ05
50	32	61.5	372	16	Q8XQH7
51	32	61.5	374	17	Q27916
52	32	61.5	387	2	Q9ROM6
53	32	61.5	425	5	Q9XVK4
54	32	61.5	432	4	Q75528
55	32	61.5	432	11	Q8R0L9
56	32	61.5	439	2	Q9R9H4
57	32	61.5	454	10	Q9C941
58	32	61.5	471	11	Q8R126
59	32	61.5	484	11	Q8VD18
60	32	61.5	505	10	Q8S300
61	32	61.5	519	10	Q9SIG8
62	32	61.5	536	16	Q8YD95
63	32	61.5	573	5	Q9VWP4
64	32	61.5	637	5	Q9Y0Y6
65	32	61.5	696	16	Q8X344
66	32	61.5	701	12	Q65568
67	32	61.5	716	11	P70521
68	32	61.5	728	12	Q9ELH4
69	32	61.5	728	12	Q9DPR8
70	32	61.5	729	12	Q9B6Q0
71	32	61.5	735	10	Q49230
72	32	61.5	743	12	Q91BV8
73	32	61.5	903	16	Q8U9X3
74	32	61.5	1435	11	Q9JIZ2
75	32	61.5	1444	11	Q9JLP0

ALIGNMENTS

RESULT 1
Q46486
ID Q46486
AC Q46486
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 23.0 kDa protein (Gcra).
GN GCRA.

Q99ur5 staphylococ
Q96yh5 sulfolobus
Q9xst4 canis famil
Q57489 bacteroides
Q9pc35 xylella fas
Q00404 homo sapien
Q8ywp1 anabaena sp
Q9uee9 homo sapien
Q8523 ascaris suu
Q95nk1 plasmodium
Q9sa71 arabidopsis
Q9qtc7 marek's dis
Q9yr01 ranid herpe
Q9v8k6 drosophila
Q97gm8 streptococ
Q9ah9 streptococ
Q9any3 streptococ
Q9avk4 pisum sativ
Q97qm9 streptococ
Q9any2 streptococ
Q9ag74 streptococ
Q8p6s4 thermoanaer
Q9any1 streptococ
Q9sin2 streptomyce
Q9ubd2 homo sapien
Q8tx62 methanopyru
Q53667 mycobacteri
Q912f0 streptomyce
Q40129 lycopersico
Q9n366 caenorhabdi
Q9vmn9 drosophila
Q9znt7 arabidopsis
Q8ui95 agrobacteri
Q8xqh7 ralstonia s
Q27916 methanobact
Q9r6m6 agrobacteri
Q9xvk4 caenorhabdi
Q75528 homo sapien
Q8r0l9 mus musculu
Q9r9h4 pseudomonas
Q9c941 arabidopsis
Q8r126 mus musculu
Q8vdi18 mus musculu
Q8s300 arabidopsis
Q9sig8 arabidopsis
Q8yds5 brucella me
Q9vwp4 drosophila
Q9y0y6 drosophila
Q8x344 escherichia
Q65568 bovine herp
P70521 rattus norv
Q9elh4 meleagrid h
Q9dpr8 meleagrid h
Q9ebg0 turkey herp
Q49230 brassica ol
Q91bv8 turkey herp
Q8u9x3 agrobacteri
Q9jiz2 mus musculu
Q9jlp0 mus musculu

OS Corynebacterium xerosis, and
 OS Corynebacterium striatum.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1725, 43770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.xerosis; STRAIN=M82B;
 RX MEDLINE=96117603; PubMed=8559800;
 RA Tauch A., Kassing F., Kalinowski J., Puhler A.;
 RT "The Corynebacterium xerosis composite transposon Tn5432 consists of
 RT two identical insertion sequences, designated IS1249, flanking the
 RT erythromycin resistance gene ermCX.";
 RL Plasmid 34:119-131(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.striatum; STRAIN=M82B;
 RX MEDLINE=20194866; PubMed=10732668;
 RA Tauch A., Kriebitz S., Kalinowski J., Puhler A.;
 RT "The 51,409-bp R-plasmid pTP10 from the multidrug-resistant clinical
 RT isolate Corynebacterium striatum M82B is composed of DNA segments
 RT initially identified in soil bacteria and in plant, animal, and human
 RT pathogens.";
 RL Mol. Gen. Genet. 263:1-11(2000).
 DR EMBL; U21300; AAC95478.1; -.
 DR EMBL; AF024666; AAG03390.1; -.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 208 AA; 23012 MW; F1504BE1ECDB85A6 CRC64;
 Query Match 59.28; Score 36; DB 2; Length 208;
 Best Local Similarity 50.0%; Pred. No. 8.6;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXHY 11
 Db 130 DVIPEKHVA 139
 :|:|:|:|:
 RESULT 2
 Q40479 PRELIMINARY; PRT; 233 AA.
 ID Q40479
 AC Q40479
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE EREP-2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BY4; TISSUE=LEAF;
 RX MEDLINE=95276459; PubMed=77556828;
 RA Ohme-Takagi M., Shinshi H.;
 RT "Ethylene-inducible DNA binding proteins that interact with an
 RT ethylene responsive element.";
 RL Plant Cell 7:173-182(1995).
 DR EMBL; D38126; BAA07324.1; -.
 DR HSP; O80337; 2GCC.
 DR TRANSFAC; T02654; -.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETRHSPLEMT.
 DR ProDom; PD001423; TF_AP2; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 233 AA; 25563 MW; 6CD16783582C0CB5 CRC64;
 Query Match 67.3%; Score 35; DB 10; Length 233;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 130 DVIPEKHVA 139
 :|:|:|:|:
 RESULT 3
 Q9LW50 PRELIMINARY; PRT; 237 AA.
 ID Q9LW50
 AC Q9LW50
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Ethylene-responsive element binding factor.
 GN NSERF2.
 OS Nicotiana sylvestris (Wood tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20399450; PubMed=10945353;
 RA Katsujima S., Koyama T., Ohme-Takagi M., Shinshi H., Sato F.;
 RT "Characterization of gene expression of NSERFs, transcription factors
 RT of basic PR genes from Nicotiana sylvestris.";
 RL Plant Cell Physiol. 41:817-824(2000).
 DR EMBL; AB016264; BAA97122.1; -.
 DR HSP; O80337; 2GCC.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETRHSPLEMT.
 DR ProDom; PD001423; TF_AP2; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 237 AA; 26243 MW; 01BC3EEB51E46298 CRC64;
 Query Match 67.3%; Score 35; DB 10; Length 237;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EVVPXGXHY 10
 Db 94 QAVVPKGRHY 103
 :|:|:|:|:
 RESULT 4
 Q98HU6 PRELIMINARY; PRT; 285 AA.
 ID Q98HU6
 AC Q98HU6
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Cytochrome c1.
 GN MLL2705.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003000; BAB49770.1; -.
 DR InterPro; IPR000345; CytC_heme_bind.
 DR InterPro; IPR002326; Cyt_C1.

```

DR Pfam: PF02167; Cytochrome_C1; 1.
DR PRINTS; PRO0603; CYTOCHROME_C1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 285 AA; 30961 MW; 31D9CDE2711747EE CRC64;

Query Match 67.3%; Score 35; DB 16; Length 285;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 11
|:| | | |
Db 194 VPEGTHY 202

RESULT 5
Q9X2E2 PRELIMINARY; PRT; 308 AA.
AC Q9X2E2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FTSH protease activity modulator HFLK.
GN TM1822.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;

RN SEQUENCE FROM N.A.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001819; AAD36885.1; -.
DR TIGR; TM1822; -.
DR InterPro; IPR001107; Band_7.
DR InterPro; IPR003130; GED.
DR InterPro; IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SM00302; GED; 1.
DR SMART; SM00244; PHB; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 308 AA; 34778 MW; ADE03603E5101A9D CRC64;

Query Match 67.3%; Score 35; DB 16; Length 308;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
|:| | | |
Db 41 VVPSGIHY 48

RESULT 6
Q38317 PRELIMINARY; PRT; 317 AA.
AC Q38317;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Lysin.
GN Lys.
OS Lactobacillus bacteriophage phi adh.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=12417;

RN SEQUENCE FROM N.A.
RP Altermann E.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[1]
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99384014; PubMed=10452953;
RA Altermann E., Klein J., Henrich B.;
RT "Primary structure and features of the genome of the Lactobacillus
gasserii temperate bacteriophage phi-adh.";
RL Gene 236:333-346(1999).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=95138034; PubMed=7836307;
RA Henrich B., Binshofer B., Blaesi U.;
RT "Primary structure and functional analysis of the lysis genes of
Lactobacillus gasserii bacteriophage phi-adh.";
RL J. Bacteriol. 177:723-732(1995).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=93231538; PubMed=8472961;
RA Fremaux C., De Antoni G., Raya R., Klaenhammer T.;
RT "Genetic organization and sequence of the region encoding integrative
functions from Lactobacillus gasserii temperate bacteriophage phi-
adh.";
RL Gene 126:61-66(1993).
[5]
RN SEQUENCE FROM N.A.
RP Engel G., Altermann E., Klein J., Henrich B.;
RT "Structure of a genome region of the Lactobacillus gasserii temperate
phage phi adh covering a repressor gene and cognate promoters.";
RL Gene 210:67-70(1998).
DR EMBL; AJ131519; CAB52540.1; -.
DR InterPro; IPR002053; GH_25.
DR InterPro; IPR003646; SH3_bac.
DR Pfam; PF01183; Glyco_hydro_25; 1.
DR ProDom; PD004620; GH_25; 1.
DR SMART; SM00287; SH3b; 1.
SQ SEQUENCE 317 AA; 34703 MW; 9FF2715EB43561C7 CRC64;

Query Match 67.3%; Score 35; DB 9; Length 317;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 11
|:| | | |
Db 60 VVPMGYHY 68

RESULT 7
Q27679 PRELIMINARY; PRT; 360 AA.
AC Q27679;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cell division protein.
GN MTH1642.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
McDougal S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;

```

RT Complete genome sequence of Methanobacterium thermoautotrophicum
 RL deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000923; AAB86115.1; -
 DR InterPro; IPR005140; eRF1_1;
 DR InterPro; IPR005141; eRF1_2;
 DR InterPro; IPR005142; eRF1_3;
 DR InterPro; IPR004405; PeIA.
 DR Pfam; PF03463; eRF1_1; 1;
 DR Pfam; PF03464; eRF1_2; 1;
 DR Pfam; PF03465; eRF1_3; 1;
 DR TIGRfams; TIGR01111; peIA; 1;
 KW Cell division; Complete proteome.
 SQ SEQUENCE 360 AA; 40814 MW; 2A000CB4B3CEFA69 CRC64;

Query Match 67.3%; Score 35; DB 17; Length 360;
 Best Local Similarity 45.5%; Pred. No. 26;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEWVPGXHY 11
 Db 98 EDLVPMSGSHHT 108
 I:::| | | |

RESULT 8

Q9E1X6 PRELIMINARY; PRT; 678 AA.
 AC Q9E1X6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 75.9 kba protein.
 OS Cercopithecine herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=35245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF275348; AAG27217.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 678 AA; 75850 MW; A17B09E30512FE3C CRC64;

Query Match 67.3%; Score 35; DB 12; Length 678;
 Best Local Similarity 50.0%; Pred. No. 52;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEWVPGXHY 10
 Db 147 EEIIPKTRY 156
 I:::| | | |

RESULT 9

Q8YJ11 PRELIMINARY; PRT; 1028 AA.
 AC Q8YJ11;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ATP-dependent DNA helicase.
 GN BMEI0275.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE-2002109; PubMed-1175668;
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
 Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,

RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AE009470; AAL51457.1; -
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; helicase_C; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 1028 AA; 112996 MW; A752B704257E219 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 1028;
 Best Local Similarity 54.5%; Pred. No. 81;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEWVPGXHY 11
 Db 76 EKIVPPGARYS 86
 I:::| | | |

RESULT 10

Q8RG86 PRELIMINARY; PRT; 1063 AA.
 AC Q8RG86;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
 GN FN0422
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 25586;
 RX MEDLINE-21886394; PubMed-11889109;
 RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 Shaleva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010554; AAL94625.1; -
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 84;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EEWVPGXHY 11
 Db 195 EIVPGLNYS 204
 I:::| | | |

RESULT 11

Q9RXN9 PRELIMINARY; PRT; 279 AA.
 AC Q9RXN9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein DR0271.
 GN DR0271.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001889; AAF09867.1; -;
 DR TIGR: DR0271; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 279 AA; 31140 MW; DCEA100E0AEB831 CRC64;
 Query Match 65.4%; Score 34; DB 16; Length 279;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 VPGXHY 11
 Db 100 VPLGRHYS 107
 RESULT 12
 Q9HLH8
 ID Q9HLH8 PRELIMINARY; PRT; 322 AA.
 AC Q9HLH8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Glucose-fructose oxidoreductase related protein.
 GN TA0250.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatiales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 RT acidophilum";
 RL Nature 407:508-513(2000).
 DR EMBL: AL445063; CAC11395.1; -;
 DR InterPro: IPR00683; GFO_IDH_MocA.
 DR Pfam: PF01408; GFO_IDH_MocA; 1.
 KW Complete proteome.
 SQ SEQUENCE 322 AA; 36918 MW; B8C239E71009D167 CRC64;
 Query Match 65.4%; Score 34; DB 17; Length 322;
 Best Local Similarity 75.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGHY 10
 Db 66 VVPGLHY 73
 RESULT 13
 O16912
 ID O16912 PRELIMINARY; PRT; 495 AA.
 AC O16912;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE F10D2.3 protein.
 GN F10D2.3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Graves T., Wohldmann P., Gillam B.;
 RT "The sequence of C. elegans cosmid F10D2.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF022972; AAC48234.1; -;
 DR InterPro: IPR004151; Sre.
 DR Pfam: PF03125; Sre; 1.
 DR SEQUENCE 495 AA; 58190 MW; 0C61139C138EE4C CRC64;
 Query Match 65.4%; Score 34; DB 5; Length 495;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGHY 10
 Db 218 ENIVPTGKHH 227
 RESULT 14
 O29966
 ID O29966 PRELIMINARY; PRT; 534 AA.
 AC O29966;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Sarcosine oxidase, subunit alpha (SOXA).
 GN AF0273.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason P.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-

```

RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -I- COFACTOR: FAD (BY SIMILARITY).
CC -I- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
DR EMBL: AE01086; AAB90959.1; -.
DR TIGR: AF0273; -.
DR InterPro: IPR000564; 2Fe2S_ferredoxin.
DR InterPro: IPR001327; FAD_pyr_redox.
DR InterPro: IPR001041; Ferredoxin.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF00111; ferr2; 1.
DR Pfam: PF00070; pyr_redox; 1.
DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
KW FAD; Flavoprotein; Hypothetical protein; Iron-sulfur; Oxidoreductase;
KW Complete proteome.
SQ SEQUENCE 534 AA; 59352 MW; A9DB03174F95093F CRC64;

Query Match 65.4%; Score 34; DB 17; Length 534;
Best Local Similarity 45.5%; Pred. No. 66;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGXHYS 11
Db 119 DKVPHGSHVT 129
::: || |||

RESULT 15
Q9ZHG7 PRELIMINARY; PRT; 822 AA.
AC Q9ZHG7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 92.4 kDa protein.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R268;
RX MEDLINE=99115568; PubMed=9916102;
RA Spellerberg B., Rozdzinski E., Martin S., Weber-Heymann J.,
RA Schnitzler N., Luetticken R., Podbielski A.;
RT "Lmb, a protein with similarities to the Irai adhesin family, mediates
RT attachment of Streptococcus agalactiae to human laminin.";
RL Infect. Immun. 67:871-878(1999).
DR EMBL: AF062533; AAD13797.1; -.
KW Hypothetical protein.
SQ SEQUENCE 822 AA; 92386 MW; 80E4EDF313481F98 CRC64;

Query Match 65.4%; Score 34; DB 2; Length 822;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
Db 350 VVPXGXHY 357
||| |||

RESULT 16
Q9XV4 PRELIMINARY; PRT; 825 AA.
AC Q9XV4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein SpY2006.
GN SpY2006.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006623; AAK34688.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 825 AA; 92649 MW; 57DF50969E6D50F4 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 825;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHY 10
Db 352 VVPXGXHY 359
||| |||

RESULT 17
Q99UR5 PRELIMINARY; PRT; 1057 AA.
AC Q99UR5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Carbamoyl-phosphate synthase large chain.
GN PYRAB OR SAVI203 OR SAI046.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Quach A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Masumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003361; BAB57365.1; -.
DR EMBL: AP003132; BAB42298.1; -.
DR HSP: P00968; 1CS0.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR004362; MGS_like.
DR InterPro: IPR001169; SHprot_acsite.
DR Pfam: PF00289; CPSase_L_chain; 2.
DR Pfam: PF02786; CPSase_L_D2; 2.
DR Pfam: PF02787; CPSase_L_D3; 1.
DR Pfam: PF02142; MGS; 1.
DR PRINTS: PR00098; CPSASE.
DR PROSITE: PS00866; CPSASE_1; 2.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_2.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 1057;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXHYS 11

```

Db 190 EIVSNGLHYS 199
 ||| | |||

RESULT 18

Q96YH5 PRELIMINARY; PRT; 1442 AA.
 ID Q96YH5
 AC Q96YH5
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Hypothetical protein S2195.
 GN S2195.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 NCBI_TaxID=111955;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kwarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7";
 RL DNA Res. 8:123-140(2001).
 KW EMBL; AP000989; BAB67302.1; -
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1442 AA; 156497 MW; D63EC2C35228121F CRC64;

Query Match 65.4%; Score 34; DB 17; Length 1442;

Best Local Similarity 50.0%; Pred. No. 2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
 ||| | |||

Db 863 EEITPTGANY 872

RESULT 19

Q9XST4 PRELIMINARY; PRT; 78 AA.
 ID Q9XST4
 AC Q9XST4
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE P97 homologous protein (Fragment).
 GN P97.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_TaxID=9615;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=THYROID;
 RC MEDLINE=20422104; PubMed=10964405;
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
 RA Christophe D.;
 RT "A method for the large-scale cloning of nuclear proteins and nuclear
 RT targeting sequences on a functional basis";
 RL Anal. Biochem. 284:231-239(2000).
 DR EMBL; AJ388531; CAB46833.1; -
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8895 MW; B62486313555FBAL CRC64;

Query Match 63.5%; Score 33; DB 6; Length 78;

Best Local Similarity 54.5%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
 ||| | |||

Db 16 EDYVPSGGEYS 26

RESULT 20

Q57489 PRELIMINARY; PRT; 139 AA.
 ID Q57489
 AC Q57489
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE DNA ligase (Fragment).
 GN DNA ligase (Fragment).
 OS Bacteroides nodosus (Dichelobacter nodosus).
 OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
 OC Dichelobacter.
 NCBI_TaxID=870;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96020672; PubMed=7476204;
 RX Moses E.K., Good R.T., Sinistaj M., Billington S.J., Langford C.J.,
 RA Rood J.I.;
 RT "A multiple site-specific DNA-inversion model for the control of Omp1
 RT phase and antigenic variation in Dichelobacter nodosus";
 RL Mol. Microbiol. 17:183-196(1995).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96257263; PubMed=8654969;
 RX Billington S.J., Sinistaj M., Cheetham B.F., Ayres A., Moses E.K.,
 RA Katz M.E., Rood J.I.;
 RT "Identifications of a native Dichelobacter nodosus plasmid and
 RT implications for the evolution of the vap regions";
 RL Gene 172:111-116(1996).
 DR EMBL; U02462; AAB12366.1; -
 DR InterPro; IPR001357; BRCT.
 DR Pfam; PF00533; BRCT; 1.
 DR SMART; SM00292; BRCT; 1.
 DR PROSITE; PS50172; BRCT; 1.
 KW Ligase.
 FT NON_TER 139
 SQ SEQUENCE 139 AA; 15052 MW; E0E110AA4B7D4708 CRC64;

Query Match 63.5%; Score 33; DB 2; Length 139;

Best Local Similarity 55.6%; Pred. No. 25;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXHYS 11
 ||| | |||

Db 21 IVPAGVHWS 29

RESULT 21

Q9PC35 PRELIMINARY; PRT; 156 AA.
 ID Q9PC35
 AC Q9PC35
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein Xfi950.
 GN Xfi950.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 NCBI_TaxID=2371;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=9A5C;
 RC MEDLINE=20365717; PubMed=10910347;
 RX Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo L.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

```

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsal S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004014; AAF84752.1; -.
DR InterPro: IPR002545; Chw.
DR SMART; SM00260; Chw; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 156 AA; 17144 MW; D8358619C6671A5D CRC64;

Query Match 63.5%; Score 33; DB 16; Length 156;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXH 9
Db 119 EEILPQGVH 127

RESULT 22
O00404 PRELIMINARY; PRT; 217 AA.
AC O00404;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04; Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P37 homologous protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=97160586; PubMed=9006920;
RA Nobukuni T., Kobayashi M., Omori A., Ichinose S., Iwanaga T.,
RA Takahashi I., Hashimoto K., Hattori S., Kaibuchi K., Miyata Y.;
RT "An Alu-linked repetitive sequence corresponding to 280 amino acids is
RT expressed in a novel bovine protein, but not in its human homologue.";
RL J. Biol. Chem. 272:2801-2807(1997).
DR EMBL; D85939; BAA20069.1; -.
SQ SEQUENCE 217 AA; 24061 MW; E404BA2E35497828 CRC64;

Query Match 63.5%; Score 33; DB 4; Length 217;
Best Local Similarity 54.5%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
Db 16 EDVPSGGEYS 26

RESULT 23
Q8YWP1 PRELIMINARY; PRT; 252 AA.
AC Q8YWP1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein ALR1563.
GN ALR1563.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kato T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003586; BAB77929.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 252 AA; 28831 MW; 925572DA5D1CA519 CRC64;

Query Match 63.5%; Score 33; DB 16; Length 252;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXHY 10
Db 235 EMIVPAGLHF 244

RESULT 24
Q9UEE9 PRELIMINARY; PRT; 299 AA.
AC Q9UEE9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BCNT protein (CRANIOFACIAL development protein 1).
GN BCNT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98267221; PubMed=9602175;
RA Takahashi I., Nobukuni T., Ohmori H., Kobayashi M., Tanaka S.,
RA Ohshima K., Okada N., Masui T., Hashimoto K., Iwashita S.;
RT "Existence of a bovine LINE repetitive insert that appears in the cDNA
RT of bovine protein BCNT in ruminant, but not in human, genomes.";
RL Gene 211:387-394(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX Strausberg R.;
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB009285; BAA31867.1; -.
DR EMBL; BC000991; AAH00991.1; -.
SQ SEQUENCE 299 AA; 33593 MW; F4A9E28B669451A CRC64;

Query Match 63.5%; Score 33; DB 4; Length 299;
Best Local Similarity 54.5%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXHYS 11
Db 16 EDVPSGGEYS 26

```

```

RESULT 25
Q08523
ID Q08523 PRELIMINARY; PRT; 412 AA.
AC Q08523; 1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2-methyl branched-chain enoyl CoA reductase precursor.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RX MEDLINE=94042988; PubMed=7693666;
RA Duran E., Komuniecki R.W., Komuniecki P.R., Wheelock M.J.,
RA Klingbeil M.M., Ma Y.C., Johnson K.R.;
RT "Characterization of cDNA clones for the 2-methyl branched-chain
RT enoyl-CoA reductase. An enzyme involved in branched-chain fatty acid
RT synthesis in anaerobic mitochondria of the parasitic nematode Ascaris
RT suum.";
RL J. Biol. Chem. 268:22391-22396(1993).
DR EMBL; L22165; AAAL6096.1; -.
DR HSP; P11310; 1EGE.
DR InterPro: IPR01552; Acyl-CoA_dh.
DR Pfam: PF00441; Acyl-CoA_dh; 1.
DR Pfam: PF02770; Acyl-CoA_dh_M; 1.
DR Pfam: PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE: PS00073; ACYL_COA_DH_2; 1.
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 16 MITOCHONDRION.
FT CHAIN 17 412 2-METHYL BRANCHED-CHAIN ENOYL COA
FT REDUCTASE
SQ SEQUENCE 412 AA; 44849 MW; 31173867DD6E0C40 CRC64;

Query Match 63.5%; Score 33; DB 5; Length 412;
Best Local Similarity 50.0%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXHY 10
| | | | |
Db 47 EEVIPKAAHH 56

```

Search completed: June 10, 2003, 13:46:37
 Job time : 27.7857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds
(without alignments)
46.744 Million cell updates/sec

Title: US-09-909-164-45
Perfect score: 50
Sequence: 1 EEVVPXGXDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	94.0	11	23	ABB80547 Hepatitis C virus
2	47	94.0	11	23	ABB80548 Hepatitis C virus
3	47	94.0	11	23	ABB80551 Hepatitis C virus
4	47	94.0	11	23	ABB80556 Hepatitis C virus
5	47	94.0	11	23	ABB80557 Hepatitis C virus
6	46	92.0	11	23	ABB80524 Hepatitis C virus
7	46	92.0	11	23	ABB80528 Hepatitis C virus
8	46	92.0	11	23	ABB80529 Hepatitis C virus
9	46	92.0	11	23	ABB80533 Hepatitis C virus
10	46	92.0	11	23	ABB80534 Hepatitis C virus

11	46	92.0	11	23	ABB80538 Hepatitis C virus
12	46	92.0	11	23	ABB80542 Hepatitis C virus
13	46	92.0	11	23	ABB80543 Hepatitis C virus
14	46	92.0	11	23	ABB80561 Hepatitis C virus
15	46	92.0	11	23	ABB80562 Hepatitis C virus
16	41	82.0	11	23	ABB80544 Hepatitis C virus
17	41	82.0	11	23	ABB80545 Hepatitis C virus
18	41	82.0	11	23	ABB80549 Hepatitis C virus
19	41	82.0	11	23	ABB80552 Hepatitis C virus
20	41	82.0	11	23	ABB80553 Hepatitis C virus
21	40	80.0	11	23	ABB80521 Hepatitis C virus
22	40	80.0	11	23	ABB80522 Hepatitis C virus
23	40	80.0	11	23	ABB80525 Hepatitis C virus
24	40	80.0	11	23	ABB80526 Hepatitis C virus
25	40	80.0	11	23	ABB80530 Hepatitis C virus
26	40	80.0	11	23	ABB80535 Hepatitis C virus
27	40	80.0	11	23	ABB80536 Hepatitis C virus
28	40	80.0	11	23	ABB80539 Hepatitis C virus
29	40	80.0	11	23	ABB80540 Hepatitis C virus
30	40	80.0	11	23	ABB80546 Hepatitis C virus
31	40	80.0	11	23	ABB80550 Hepatitis C virus
32	40	80.0	11	23	ABB80554 Hepatitis C virus
33	40	80.0	11	23	ABB80555 Hepatitis C virus
34	40	80.0	11	23	ABB80559 Hepatitis C virus
35	40	80.0	11	23	ABB80563 Hepatitis C virus
36	40	80.0	11	23	ABB80564 Hepatitis C virus
37	40	80.0	11	23	ABB80565 Hepatitis C virus
38	40	80.0	11	23	ABB80566 Hepatitis C virus
39	40	80.0	11	23	ABB80567 Hepatitis C virus
40	40	80.0	11	23	ABB80568 Hepatitis C virus
41	39	78.0	11	23	ABB80523 Hepatitis C virus
42	39	78.0	11	23	ABB80527 Hepatitis C virus
43	39	78.0	11	23	ABB80531 Hepatitis C virus
44	39	78.0	11	23	ABB80532 Hepatitis C virus
45	39	78.0	11	23	ABB80537 Hepatitis C virus
46	39	78.0	11	23	ABB80541 Hepatitis C virus
47	39	78.0	11	23	ABB80558 Hepatitis C virus
48	39	78.0	11	23	ABB80560 Hepatitis C virus
49	35	70.0	143	22	ABG60256 Drosophila melanog
50	35	70.0	150	21	ACG28379 Arabidopsis thalia
51	35	70.0	299	22	ABG00168 Novel human diagno
52	35	70.0	587	21	AAV74287 Drosophila melanog
53	35	70.0	692	22	ABG64835 Bovine mammary tis
54	34	68.0	150	22	AAAB7623 Arabidopsis thalia
55	34	68.0	175	21	ACG10069 C glutamicum prote
56	34	68.0	283	22	ACG91127 Novel central nerv
57	34	68.0	576	22	AAU87272 Novel human diagno
58	34	68.0	1022	22	ABG03621 Novel human diagno
59	34	68.0	1022	22	ABG05826 Novel human diagno
60	34	68.0	1022	22	ABG08173 Drosophila melanog
61	34	68.0	2778	22	ABG58683 Eubacterial DNA po
62	33	66.0	25	23	ABG52372 MAb L243 VH region
63	33	66.0	140	16	AAAG4232 Humanized antibody
64	33	66.0	140	16	AAAG4235 CDR-grafted L243-g
65	33	66.0	140	16	AAAG4265 MHC-II MAb L243 he
66	33	66.0	140	16	AAAG4257 Listeria monocytog
67	33	66.0	165	23	ABBA4059 Group B Streptococ
68	33	66.0	222	22	AAU03629 Streptococcus poly
69	33	66.0	222	23	ABP26468 Murine JNK3 bindin
70	33	66.0	244	21	AAAB12881 Murine JNK3 bindin
71	33	66.0	484	21	AAAB12882 Staphylococcus epi
72	33	66.0	563	23	ABP27069 Propionibacterium
73	32	64.0	139	23	ABP40279 Drosophila melanog
74	32	64.0	150	22	AAU43140
75	32	64.0	275	22	ABG63063

ALIGNMENTS

RESULT 1
ABB80547
ID ABB80547 standard; peptide; 11 AA.

FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 94.0%; Score 47; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0081;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDXYS 11
 ||||| |||
 Db 1 EEVVPXGXSDYS 11

RESULT 4

ID ABB80556 standard; peptide; 11 AA.

XX ABB80556;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 94.0%; Score 47; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0081;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDXYS 11
 ||||| |||
 Db 1 EEVVPXGXSDYS 11

RESULT 5

ID ABB80557 standard; peptide; 11 AA.

XX ABB80557;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ
 Query Match 94.0%; Score 47; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0081;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db ||||| ||||
 1 EEVVPXGSDYS 11
 RESULT 6
 ABB80524
 ID ABB80524 standard; peptide; 11 AA.
 XX AC ABB80524;
 XX DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT W0200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ
 Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db ||||| ||||
 1 EEVVPXGMDYS 11
 RESULT 7
 ABB80528
 ID ABB80528 standard; peptide; 11 AA.
 XX AC ABB80528;
 XX DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT W0200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 ID ABB80529
 XX 1 EEVVPXGMDYS 11

RESULT 8

ABB80529
 ID ABB80529 standard; peptide; 11 AA.

XX AC ABB80529;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory
 activity useful for treating disorders associated with hepatitis C
 virus protease

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 invention are alpha-ketoamide peptide analogues. The peptides have
 virucide activity, and are useful for treating and in the manufacture of
 a medicament to treat disorders associated with HCV protease. A
 pharmaceutical composition comprising the peptide as an active ingredient
 is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 ID ABB80533
 XX 1 EEVVPXGMDYS 11

RESULT 9

ABB80533
 ID ABB80533 standard; peptide; 11 AA.

XX AC ABB80533;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #13.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory
 activity useful for treating disorders associated with hepatitis C
 virus protease

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 invention are alpha-ketoamide peptide analogues. The peptides have
 virucide activity, and are useful for treating and in the manufacture of
 a medicament to treat disorders associated with HCV protease. A
 pharmaceutical composition comprising the peptide as an active ingredient
 is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 ID ABB80534
 XX 1 EEVVPXGMDYS 11

RESULT 10

ABB80534

ID XX ABB80534 standard; peptide; 11 AA.
 AC ABB80534;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #14.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT
 FT Misc-difference 9
 FT Modified-site 11
 FT /note= "D-form residue"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGDYS 11
 Db 1 EEVVPXGDYS 11
 RESULT 11
 ABB80538
 ID ABB80538 standard; peptide; 11 AA.
 XX
 AC ABB80538;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
 XX

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT
 FT Misc-difference 9
 FT Modified-site 11
 FT /note= "D-form residue"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGDYS 11
 Db 1 EEVVPXGDYS 11
 RESULT 12
 ABB80542
 ID ABB80542 standard; peptide; 11 AA.
 XX
 AC ABB80542;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
PD
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
PI
XX WPI; 2002-361643/39.
DR
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
PT
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.013; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGQDYS 11
RESULT 13
ABB80543
ID ABB80543 standard; peptide; 11 AA.
XX
XX ABB80543;
AC
XX 08-OCT-2002 (first entry)
DT
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #23.
DE
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "D-form residue"
FT

FT /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
PD
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
PA
XX Lim-wilby M, Levy OE, Brunck TK;
PI
XX WPI; 2002-361643/39.
DR
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
PT
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.013; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGQDYS 11
RESULT 14
ABB80561
ID ABB80561 standard; peptide; 11 AA.
XX
XX ABB80561;
AC
XX 08-OCT-2002 (first entry)
DT
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.
DE
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 8 /note= "Oxymethionine"
FT Modified-site 11 /note= "C-terminal amide"
FT
XX WO200208251-A2.
PN
XX 31-JAN-2002.
PD
XX

PF 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. NO. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDXYS 11
 DB ||||| |||
 1 EEVVPXGMDYS 11

RESULT 15
 ABB80562
 ID ABB80562 standard; peptide; 11 AA.
 AC ABB80562;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.

PA (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 92.0%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. NO. 0.013;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDXYS 11
 DB ||||| |||
 1 EEVVPXGMDYS 11

RESULT 16
 ABB80544
 ID ABB80544 standard; peptide; 11 AA.
 XX ABB80544;
 AC 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX

PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGXGDSYS 11
Db 1 EEVVPXGXGTSYS 11

RESULT 17

ABB80545
ID ABB80545 standard; peptide; 11 AA.

XX AC ABB80545;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #25.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGXGDSYS 11
Db 1 EEVVPXGXGTSYS 11

RESULT 18

ABB80549
ID ABB80549 standard; peptide; 11 AA.

XX AC ABB80549;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #29.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 PS Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 80.0%; Score 40; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDS 11
 ||||| ||
 DB 1 EEVVPXGMSYS 11
 RESULT 22
 ABB80522
 ID ABB80522 standard; peptide; 11 AA.
 XX
 AC ABB80522;
 XX
 DT 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT WO200208251-A2.

FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 PS Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 80.0%; Score 40; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDS 11
 ||||| ||
 DB 1 EEVVPXGMSYS 11
 RESULT 23
 ABB80525
 ID ABB80525 standard; peptide; 11 AA.
 XX
 AC ABB80525;
 XX
 DT 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN

XX PD 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX PA Lim-wilby M, Levy OE, Brunck TK;
XX PI WPI; 2002-361643/39.
XX DR
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease -
XX
XX Claim 17; Page 64; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
Query Match 80.0%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGMSYS 11
RESULT 24
ABB80526
ID ABB80526 standard; peptide; 11 AA.
XX
XX ABB80526;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX residue 7"
XX
XX Misc-difference 8 /note= "D-form residue"
XX Misc-difference 9 /note= "D-form residue"
XX Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease -
XX
XX Claim 17; Page 64; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
Query Match 80.0%; Score 40; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGMSYS 11
RESULT 25
ABB80530
ID ABB80530 standard; peptide; 11 AA.
XX
XX ABB80530;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #10.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease -

XX
PS
XX
XX

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX

SQ Sequence 11 AA;

Query Match 80.0%; Score 40; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.2;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 EEVVPXGXYS 11

|||||

Db

1 EEVVPXGGSYS 11

Search completed: June 10, 2003, 13:39:11

Job time : 31.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds
(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-45
Perfect score: 50
Sequence: 1 BEVVPXGXDS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

Issued Patents.AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	66.0	140	4	US-08-569-147-76
2	33	66.0	140	4	US-08-569-147-82
3	32	64.0	139	4	US-09-134-001C-5124
4	31	62.0	622	2	US-08-459-146-2
5	31	62.0	622	2	US-08-459-065-2
6	31	62.0	739	4	US-09-413-814-86
7	30	60.0	59	4	US-08-963-851-14
8	30	60.0	121	4	US-09-152-060-68
9	30	60.0	121	4	US-09-152-060-85
10	30	60.0	122	2	US-08-879-995A-1
11	30	60.0	122	3	US-09-215-096-1
12	30	60.0	231	3	US-08-926-842B-20
13	30	60.0	241	3	US-08-926-842B-21
14	30	60.0	241	3	US-08-834-776A-2
15	30	60.0	341	4	US-08-853-948B-4
16	30	60.0	348	4	US-08-853-948B-5
17	30	60.0	368	4	US-09-697-367-24
18	30	60.0	421	4	US-09-093-448-4
19	30	60.0	478	2	US-09-040-799-3
20	30	60.0	478	4	US-09-093-448-1
21	30	60.0	478	4	US-09-093-448-2
22	30	60.0	478	4	US-09-093-448-3
23	30	60.0	1068	2	US-08-429-054A-11
24	30	60.0	1068	2	US-08-718-777-7
25	30	60.0	1068	3	US-09-051-341-7
26	29	58.0	100	2	US-09-047-125-27
27	29	58.0	100	3	US-07-736-335E-27

ALIGNMENTS

RESULT 1

US-08-569-147-76

; Sequence 76, Application US/08569147

; Patent No. 6180377

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: HUMANISED ANTIBODIES

; NUMBER OF SEQUENCES: 95

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESS: No. 6180377ris, LLP

; CITY: Philadelphia

; STATE: PA

; COUNTRY: U.S.A.

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

Sequence 3003, Ap
Sequence 10, Appl
Sequence 4, Appl
Sequence 14, Appl
Sequence 4, Appl
Sequence 9, Appl
Sequence 5, Appl
Sequence 236, App
Sequence 236, App
Sequence 62, Appl
Sequence 62, Appl
Sequence 3897, Ap
Sequence 13, Appl
Sequence 2, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 3973, Ap
Sequence 4, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 403, App
Sequence 3, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 6, Appl

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/569,147
;; FILING DATE: 25-March-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trujillo, Doreen Yanko
;; REGISTRATION NUMBER: 35,719
;; REFERENCE/DOCKET NUMBER: CARP-0047
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 76:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 140 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-569-147-76

Query Match 66.0%; Score 33; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDY 10
Db 122 VVPTGFDY 129

RESULT 2
US-08-569-147-82
;; Sequence 82, Application US/08569147
;; Patent No. 6180377
;; GENERAL INFORMATION:
;; APPLICANT: HUMANISED ANTIBODIES
;; TITLE OF INVENTION: HUMANISED ANTIBODIES
;; NUMBER OF SEQUENCES: 95
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
;; ADDRESSEE: No. 6180377ris, LLP
;; STREET: One Liberty Place - 46th Floor
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: U.S.A.
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/569,147
;; FILING DATE: 25-March-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trujillo, Doreen Yanko
;; REGISTRATION NUMBER: 35,719
;; REFERENCE/DOCKET NUMBER: CARP-0047
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 568-3100
;; TELEFAX: (215) 568-3439
;; INFORMATION FOR SEQ ID NO: 82:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 140 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-569-147-82

Query Match 66.0%; Score 33; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDY 10
Db 122 VVPTGFDY 129

RESULT 3
US-09-134-001C-5124
;; Sequence 5124, Application US/09134001C
;; Patent No. 6380370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 5124
;; LENGTH: 139
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5124

Query Match 64.0%; Score 32; DB 4; Length 139;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDYS 11
Db 32 IVPFGHDYN 40

RESULT 4
US-08-459-146-2
;; Sequence 2, Application US/08459146
;; Patent No. 5866405
;; GENERAL INFORMATION:
;; APPLICANT: Choi, Gil Ho
;; APPLICANT: Nuss, Donald Lee
;; TITLE OF INVENTION: Genetically Engineered Transmissible
;; TITLE OF INVENTION: Hypovirulence
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
;; STREET: 340 Kingsland Street
;; CITY: Nutley
;; STATE: New Jersey
;; COUNTRY: U.S.A.
;; ZIP: 07110
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/459,146
;; FILING DATE: 02-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/832,117
;; FILING DATE: 06-FEB-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Roseman, Catherine R
;; REGISTRATION NUMBER: 34,240
;; REFERENCE/DOCKET NUMBER: 8589
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (201) 235-6208
;; TELEFAX: (201) 235-3500
;; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 622 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Endothia parasitica (Cryptonectria
STRAIN: EP713
US-08-459-146-2

Query Match 62.0%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
DB 31 EEVVPAG 37

RESULT 5
US-08-459-065-2
; Sequence 2, Application US/08459065
; Patent No. 5882642
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil Ho
; TITLE OF INVENTION: Genetically Engineered Transmissible
; TITLE OF INVENTION: Hypovirulence
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,065
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/832,117
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8589
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Endothia parasitica (Cryptonectria
; ORGANISM: parasitica)
; STRAIN: EP713
US-08-459-065-2

Query Match 62.0%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
DB 31 EEVVPAG 37

RESULT 6
US-09-413-814-86
; Sequence 86, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 86
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-86

Query Match 62.0%; Score 31; DB 4; Length 739;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGXDYS 11
| | | | |
DB 663 IPLGXDYS 670

RESULT 7
US-08-963-851-14
; Sequence 14, Application US/08963851
; Patent No. 6300116
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/08/963,851
; CURRENT FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-08-963-851-14

Query Match 60.0%; Score 30; DB 4; Length 59;
Best Local Similarity 45.5%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11

Db 38 EXHIPGLEYs 48

RESULT 8

US-09-152-060-68
; Sequence 68, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/152,060
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-152-060-68

Query Match 60.0%; Score 30; DB 4; Length 121;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

RESULT 9

US-09-152-060-85
; Sequence 85, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/152,060
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30

; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-152-060-85

Query Match 60.0%; Score 30; DB 4; Length 121;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

RESULT 10

US-08-879-995A-1
; Sequence 1, Application US/08879995A
; Patent No. 5985606
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/879,995A
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids

; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BRAITUT03
 ; CLONE: 2109906
 ; US-08-879-995A-1

Query Match 60.0%; Score 30; DB 2; Length 122;
 Best Local Similarity 85.7%; Pred. No. 55;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
 | | | | |
 Db 28 EEVPPGG 34

RESULT 11
 US-09-215-096-1
 ; Sequence 1, Application US/09215096
 ; Patent No. 6008194
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/215,096
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/879,995
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0326 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 122 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BRAITUT03
 ; CLONE: 2109906
 ; US-09-215-096-1

Query Match 60.0%; Score 30; DB 3; Length 122;
 Best Local Similarity 85.7%; Pred. No. 55;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
 | | | | |
 Db 28 EEVPPGG 34

RESULT 12

US-08-926-842B-20
 ; Sequence 20, Application US/08926842B
 ; Patent No. 6030807
 ; GENERAL INFORMATION:
 ; APPLICANT: Sa-No. 6030807ueira, Isabel
 ; APPLICANT: de Lencastre, Herminia
 ; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/926,842B
 ; FILING DATE: 10-SEP-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-089 N
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201 487-5800
 ; TELEFAX: 201 343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 231 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; US-08-926-842B-20

Query Match 60.0%; Score 30; DB 3; Length 231;
 Best Local Similarity 55.6%; Pred. No. 11e-02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
 : | | | |
 Db 41 IKPSGVDYS 49

RESULT 13

US-08-926-842B-21
 ; Sequence 21, Application US/08926842B
 ; Patent No. 6030807
 ; GENERAL INFORMATION:
 ; APPLICANT: Sa-No. 6030807ueira, Isabel
 ; APPLICANT: de Lencastre, Herminia
 ; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,842B
; FILING DATE: 10-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-089 N
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Salmonella typhimurium
; US-09-926-842B-21

```

```

Query Match 60.0%; Score 30; DB 3; Length 240;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 VVPXGXDXYS 11
; | | | | |
DB 41 IKPSGVDXYS 49

```

RESULT 14

```

US-08-834-776A-2
; Sequence 2, Application US/08834776A
; Patent No. 6060241
; GENERAL INFORMATION:
; APPLICANT: Cortesey-Theulaz, Irene
; TITLE OF INVENTION: Compositions and Methods Relating to
; TITLE OF INVENTION: Drug Discovery and Detection and Treatment of
; TITLE OF INVENTION: Gastrointestinal Diseases
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,776A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ashton, Nina M.
; REGISTRATION NUMBER: 37,273
; REFERENCE/DOCKET NUMBER: GAST-001/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5000
; TELEFAX: 415-857-0663
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-834-776A-2

```

```

Query Match 60.0%; Score 30; DB 3; Length 241;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXG 7
; | | | | |
DB 199 EEIVPAG 205

```

RESULT 15

```

US-08-853-948B-4
; Sequence 4, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Citrus unshiu
; FEATURE:
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asn, Asp,
; OTHER INFORMATION: Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe,
; OTHER INFORMATION: Pro, Ser, Thr, Trp, Tyr, or Val
; US-08-853-948B-4

```

```

Query Match 60.0%; Score 30; DB 4; Length 341;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 VVPXGXDXYS 11
; | | | | |
DB 228 VIPPGMDFS 236

```

RESULT 16

```

US-08-853-948B-5
; Sequence 5, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Citrus unshiu
; US-08-853-948B-5

```

```

Query Match 60.0%; Score 30; DB 4; Length 348;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 VVPXGXDXYS 11
; | | | | |

```

Db 234 VIPPGMDFS 242

RESULT 17

US-09-697-367-24
; Sequence 24, Application US/09697367
; Patent No. 6323015
; GENERAL INFORMATION:
; APPLICANT: Orozco Jr., Emil M.
; APPLICANT: Calimi, Perry G.
; APPLICANT: Weng, Zude
; APPLICANT: Tarczynski, Mitchell
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: B1166 US NA
; CURRENT APPLICATION NUMBER: US/09/697,367
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/084,529
; PRIOR FILING DATE: 1998-MAY-07
; PRIOR APPLICATION NUMBER: PCT/US99/09865
; PRIOR FILING DATE: 1999-MAY-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-09-697-367-24

Query Match 60.0%; Score 30; DB 4; Length 368;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDXS 11.
|:|:|:|:
Db 217 VIPPGMDFS 225

RESULT 18

US-09-093-448-4
; Sequence 4, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-093-448-4

Query Match 60.0%; Score 30; DB 4; Length 421;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGXDX 10
|:|:|:|:
Db 412 KEVVSXGDDY 421

RESULT 19

US-09-040-799-3
; Sequence 3, Application US/09040799
; Patent No. 5885820
; GENERAL INFORMATION:
; APPLICANT: CHANG, YIE-HWA

; TITLE OF INVENTION: CLONE OF A NUCLEOTIDE SEQUENCE ENCODING
; TITLE OF INVENTION: A PROTEIN HAVING TWO FUNCTIONS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/040,799
; APPLICATION NUMBER: US/09/040,799
; FILING DATE: 18-MAR-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 16153-4639
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-040-799-3

Query Match 60.0%; Score 30; DB 2; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGXDX 10
|:|:|:|:
Db 469 KEVVSXGDDY 478

RESULT 20

US-09-093-448-1
; Sequence 1, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-093-448-1

Query Match 60.0%; Score 30; DB 4; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGXDX 10
|:|:|:|:
Db 469 KEVVSXGDDY 478

```
RESULT 21
US-09-093-448-2
; Sequence 2, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-093-448-2

Query Match          60.0%; Score 30; DB 4; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGXDY 10
   :||| | ||
Db 469 KEVVSRGDDY 478

RESULT 22
US-09-093-448-3
; Sequence 3, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-093-448-3

Query Match          60.0%; Score 30; DB 4; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGXDY 10
   :||| | ||
Db 469 KEVVSRGDDY 478

RESULT 23
US-08-429-054A-11
; Sequence 11, Application US/08429054A
; Patent No. 5917126
; GENERAL INFORMATION:
; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,
; APPLICANT: JEAN; VOELKER, TONI; GERVASIS, MONICA
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),
; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE
; TITLE OF INVENTION: CDNA FOR MODIFYING SPS EXPRESSION IN PLANT CELLS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: BIERMAN AND MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,054A
; FILING DATE: 26-APR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 842,337
; FILING DATE: 20-March-1992
; APPLICATION NUMBER: PCT/FR 91/00593
; FILING DATE: 18-July-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: French 90402094.9
; FILING DATE: 20-July-1990
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles A. Musierlian
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068
; TYPE: Amino acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
US-08-429-054A-11

Query Match          60.0%; Score 30; DB 2; Length 1068;
Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXGDXS 11
   |:||| |||
Db 435 VIPPGMDFS 443

RESULT 24
US-08-718-777-7
; Sequence 7, Application US/08718777
; Patent No. 5981852
; GENERAL INFORMATION:
; APPLICANT: Van Assche, C.
; APPLICANT: Lando, D.
; APPLICANT: Bruneau, J. M.
; APPLICANT: Voelker, T.
; APPLICANT: Gervais, M.
; TITLE OF INVENTION: MODIFICATION OF SUCROSE
; TITLE OF INVENTION: PHOSPHATE
; TITLE OF INVENTION: SYNTHASE IN PLANTS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Barbara Rae-Venter
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,777
FILING DATE: NOT YET ASSIGNED
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,471
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE.072.02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)328-4400
TELEFAX: (415)328-4477
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-777-7

Query Match 60.0%; Score 30; DB 2; Length 1068;

Best Local Similarity 55.6%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 2;

QY 3 VVPXGXDYS 11
|:| | |:
Db 435 VIPPGMDFS 443

RESULT 25

US-09-051-341-7
Sequence 7, Application US/09051341
Patent No. 6124528
GENERAL INFORMATION:
APPLICANT: Shewmaker, C. K.
TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rae-Venter Law Group, P.C.
STREET: 260 Sheridan Avenue, Suite 440
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/051,341
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17351
FILING DATE: 25-OCT-1996
APPLICATION NUMBER: US 08/549,016
FILING DATE: 27-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/372,200
FILING DATE: 12-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter, Ph.D.,
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE.110.02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)328-4400

TELEFAX: (415)328-4477
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-051-341-7

Query Match 60.0%; Score 30; DB 3; Length 1068;
Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
|:| | |:
Db 435 VIPPGMDFS 443

Search completed: June 10, 2003, 13:51:38
Job time : 11.6429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds
(without alignments)
75.710 Million cell updates/sec

Title: US-09-909-164-45
Perfect score: 50
Sequence: 1 EEVVPXGXDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	68.0	283	9 US-09-738-626-4881	Sequence 4881, Ap
2	33	66.0	440	9 US-09-813-408-27	Sequence 27, Appl
3	32	64.0	3472	9 US-10-027-805-4	Sequence 4, Appl
4	32	64.0	3472	9 US-10-034-623-4	Sequence 4, Appl
5	32	64.0	3472	9 US-10-027-801-4	Sequence 4, Appl
6	31	62.0	299	10 US-09-815-242-10697	Sequence 10697, A
7	31	62.0	653	9 US-09-820-843A-26	Sequence 26, Appl
8	31	62.0	846	10 US-09-815-242-13904	Sequence 13904, A
9	30	60.0	7	9 US-09-909-062-1	Sequence 1, Appl
10	30	60.0	7	9 US-09-909-062-9	Sequence 9, Appl
11	30	60.0	7	9 US-09-909-062-130	Sequence 130, Appl
12	30	60.0	11	9 US-09-943-123-22	Sequence 22, Appl
13	30	60.0	59	10 US-09-948-080-14	Sequence 14, Appl
14	30	60.0	79	9 US-09-764-891-4601	Sequence 4601, Ap
15	30	60.0	121	9 US-09-852-797-68	Sequence 68, Appl
16	30	60.0	121	9 US-09-852-797-85	Sequence 85, Appl
17	30	60.0	121	10 US-09-853-161-68	Sequence 68, Appl
18	30	60.0	121	10 US-09-853-161-85	Sequence 85, Appl
19	30	60.0	121	10 US-09-852-659A-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-09-738-626-4881
; Sequence 4881, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO

20	30	60.0	121	10	US-09-852-659A-85	Sequence 85, Appl
21	30	60.0	135	9	US-09-992-598-359	Sequence 359, App
22	30	60.0	135	9	US-09-989-293A-359	Sequence 359, App
23	30	60.0	135	9	US-09-989-735-359	Sequence 359, App
24	30	60.0	135	9	US-09-990-444-359	Sequence 359, App
25	30	60.0	135	9	US-09-989-730-359	Sequence 359, App
26	30	60.0	135	9	US-09-990-436-359	Sequence 359, App
27	30	60.0	135	9	US-09-991-181-359	Sequence 359, App
28	30	60.0	135	9	US-09-993-687-359	Sequence 359, App
29	30	60.0	135	9	US-09-989-734-359	Sequence 359, App
30	30	60.0	135	9	US-09-997-653-359	Sequence 359, App
31	30	60.0	135	9	US-10-174-590-444	Sequence 444, App
32	30	60.0	135	9	US-10-176-758-444	Sequence 444, App
33	30	60.0	135	9	US-10-175-737-444	Sequence 444, App
34	30	60.0	135	9	US-09-993-667-359	Sequence 359, App
35	30	60.0	135	9	US-10-173-706-444	Sequence 444, App
36	30	60.0	135	9	US-10-175-738-444	Sequence 444, App
37	30	60.0	135	9	US-10-175-752-444	Sequence 444, App
38	30	60.0	135	9	US-10-176-482-444	Sequence 444, App
39	30	60.0	135	9	US-10-176-757-444	Sequence 444, App
40	30	60.0	135	9	US-10-176-913-444	Sequence 444, App
41	30	60.0	135	9	US-10-180-552-444	Sequence 444, App
42	30	60.0	135	9	US-10-180-557-444	Sequence 444, App
43	30	60.0	135	9	US-09-990-438-359	Sequence 359, App
44	30	60.0	135	9	US-09-990-562-359	Sequence 359, App
45	30	60.0	135	9	US-09-997-428-359	Sequence 359, App
46	30	60.0	135	9	US-09-997-666-359	Sequence 359, App
47	30	60.0	135	9	US-10-173-700-444	Sequence 444, App
48	30	60.0	135	9	US-10-174-572-444	Sequence 444, App
49	30	60.0	135	9	US-10-174-579-444	Sequence 444, App
50	30	60.0	135	9	US-10-174-582-444	Sequence 444, App
51	30	60.0	135	9	US-10-174-588-444	Sequence 444, App
52	30	60.0	135	9	US-10-175-739-444	Sequence 444, App
53	30	60.0	135	9	US-10-175-740-444	Sequence 444, App
54	30	60.0	135	9	US-10-176-488-444	Sequence 444, App
55	30	60.0	135	9	US-10-176-492-444	Sequence 444, App
56	30	60.0	135	9	US-10-176-747-444	Sequence 444, App
57	30	60.0	135	9	US-10-176-750-444	Sequence 444, App
58	30	60.0	135	9	US-10-176-985-444	Sequence 444, App
59	30	60.0	135	9	US-10-176-987-444	Sequence 444, App
60	30	60.0	135	9	US-10-176-991-444	Sequence 444, App
61	30	60.0	135	9	US-10-176-993-444	Sequence 444, App
62	30	60.0	135	9	US-10-176-993-444	Sequence 444, App
63	30	60.0	135	9	US-10-184-658-444	Sequence 444, App
64	30	60.0	135	9	US-10-227-884-108	Sequence 108, App
65	30	60.0	135	9	US-09-990-711-359	Sequence 359, App
66	30	60.0	135	9	US-10-173-695-444	Sequence 444, App
67	30	60.0	135	9	US-10-173-697-444	Sequence 444, App
68	30	60.0	135	9	US-10-173-705-444	Sequence 444, App
69	30	60.0	135	9	US-10-174-576-444	Sequence 444, App
70	30	60.0	135	9	US-10-174-585-444	Sequence 444, App
71	30	60.0	135	9	US-10-174-586-444	Sequence 444, App
72	30	60.0	135	9	US-10-175-747-444	Sequence 444, App
73	30	60.0	135	9	US-10-176-481-444	Sequence 444, App
74	30	60.0	135	9	US-10-176-485-444	Sequence 444, App
75	30	60.0	135	9	US-10-176-485-444	Sequence 444, App

; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4881
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4881

Query Match 68.0%; Score 34; DB 9; Length 283;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VPXGXDYS 11
||| |||
Db 56 VPAGADYS 63

RESULT 2

US-09-813-408-27
; Sequence 27, Application US/09813408
; Publication No. US20030049619A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Maris, Barry
; TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Lib
; TITLE OF INVENTION: Of Polynucleotides
; FILE REFERENCE: HER0041
; CURRENT APPLICATION NUMBER: US/09/813,408
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match 66.0%; Score 33; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGXDY 10
||| |||
Db 120 EVLPWGVVDY 128

RESULT 3

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCP0002A
; CURRENT APPLICATION NUMBER: US/10/027,806
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29

; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 64.0%; Score 32; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 11e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVPXGXDXYS 11
|:|:| |:
Db 2294 EDVIPRGISFS 2304

RESULT 4

US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCP0002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 64.0%; Score 32; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 11e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVPXGXDXYS 11
|:|:| |:
Db 2294 EDVIPRGISFS 2304

RESULT 5

US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCP0002A
; CURRENT APPLICATION NUMBER: US/10/027,801
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match 64.0%; Score 32; DB 9; Length 3472;

Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDS 11
|:|:|:|:|:
Db 2294 EDVIPRGISFS 2304

RESULT 6
US-09-815-242-10697
; Sequence 10697, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10697
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10697

Query Match 62.0%; Score 31; DB 10; Length 299;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10
|:|:|:|:|:
Db 218 EQITPTGIEY 227

RESULT 7
US-09-820-843A-26
; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 653
; TYPE: PRT

; ORGANISM: Vibrio cholerae
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: iron(III) ABC transporter, permease protein
; NAME/KEY: misc-feature
; OTHER INFORMATION: gi|9654609
US-09-820-843A-26

Query Match 62.0%; Score 31; DB 9; Length 653;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
|:|:|:|:|:
Db 300 EEVVPXG 306

RESULT 8
US-09-815-242-13904
; Sequence 13904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13904
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13904

Query Match 62.0%; Score 31; DB 10; Length 846;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10
|:|:|:|:|:
Db 762 EDVVDGVDY 771

RESULT 9
US-09-909-062-1
; Sequence 1, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavallabhan, Vijay M

APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLTATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-1

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 1 EEVVPXG 7

RESULT 10
US-09-909-062-9
Sequence 9, Application US/09909062
Publication No. US20030036501A1
GENERAL INFORMATION:
APPLICANT: Saksena, Anil K
APPLICANT: Girijavaliabhan, Viyyor M
APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 7
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLTATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-9

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 1 EEVVPXG 7

RESULT 11
US-09-909-062-130
Sequence 130, Application US/09909062
Publication No. US20030036501A1
GENERAL INFORMATION:
APPLICANT: Saksena, Anil K
APPLICANT: Girijavaliabhan, Viyyor M
APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 130
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLTATION
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (7)..(7)
OTHER INFORMATION: AMIDATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-130

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

Db 1 EEVVPXG 7
|||||

RESULT 12

US-09-943-123-22
; Sequence 22, Application US/09943123
; Publication No. US20020182701A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Y-H
; APPLICANT: VETRO, J.A.
; APPLICANT: MICKA, W.S.
; TITLE OF INVENTION: Dominant Negative Variants of Methionine Aminopeptidase
; FILE REFERENCE: 16153-8007
; CURRENT APPLICATION NUMBER: US/09/943,123
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-943-123-22

Query Match 60.0%; Score 30; DB 9; Length 11;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDY 10
: ||| | ||
Db 2 KEVSKGDDY 11

RESULT 13

US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102702A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAU
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match 60.0%; Score 30; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDYS 11
: : | | : ||
Db 38 EKHIPGLEYS 48

RESULT 14

US-09-764-891-4601
; Sequence 4601, Application US/09764891

; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4601
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-891-4601

Query Match 60.0%; Score 30; DB 9; Length 79;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EEVVPXGXDYS 11
: : | | |||
Db 41 ELIPGISDYS 50

RESULT 15

US-09-852-797-68
; Sequence 68, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-797-68

Query Match 60.0%; Score 30; DB 9; Length 121;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
|||||

Db 28 EEWVPGG 34

RESULT 16

US-09-852-797-85

; Sequence 85, Application US/09852797

; Patent No. US20020172994A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P2

; CURRENT APPLICATION NUMBER: US/09/852,797

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 85

; LENGTH: 121

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (67)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (89)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-852-797-85

Query Match

Best Local Similarity 60.0%; Score 30; DB 9; Length 121;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEWVPGG 7

Db 28 EEWVPGG 34

RESULT 17

US-09-853-161-68

; Sequence 68, Application US/09853161

; Patent No. US20020076756A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P3

; CURRENT APPLICATION NUMBER: US/09/853,161

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 68

; LENGTH: 121

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-853-161-68

Query Match

Best Local Similarity 60.0%; Score 30; DB 10; Length 121;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEWVPGG 7

Db 28 EEWVPGG 34

RESULT 18

US-09-853-161-85

; Sequence 85, Application US/09853161

; Patent No. US20020076756A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P3

; CURRENT APPLICATION NUMBER: US/09/853,161

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 85

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (67)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (89)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-853-161-85

Query Match

Best Local Similarity 60.0%; Score 30; DB 10; Length 121;

Mismatches 0; Conservative 0; Indels 1; Gaps 0;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

Db 28 EEVVPXG 34

RESULT 19

US-09-852-659A-68

Sequence 68, Application US/09852659A

Patent No. US20020077287A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 28 Human Secreted Proteins

FILE REFERENCE: P2003P4

CURRENT APPLICATION NUMBER: US/09/852,659A

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: 60/265,583

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/152,060

PRIOR FILING DATE: 1998-09-11

PRIOR APPLICATION NUMBER: PCT/US98/04858

PRIOR FILING DATE: 1998-03-12

PRIOR APPLICATION NUMBER: 60/040,762

PRIOR FILING DATE: 1997-03-14

PRIOR APPLICATION NUMBER: 60/040,710

PRIOR FILING DATE: 1997-03-14

PRIOR APPLICATION NUMBER: 60/050,934

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,100

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,357

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,189

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/057,765

PRIOR FILING DATE: 1997-09-05

PRIOR APPLICATION NUMBER: 60/048,970

PRIOR FILING DATE: 1997-06-06

PRIOR APPLICATION NUMBER: 60/068,368

PRIOR FILING DATE: 1997-12-19

NUMBER OF SEQ ID NOS: 121

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 68

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

US-09-852-659A-68

Query Match

Best Local Similarity 60.0%; Score 30; DB 10; Length 121;

Mismatches 0; Conservative 0; Indels 1; Gaps 0;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

Db 28 EEVVPXG 34

RESULT 20

US-09-852-659A-85

Sequence 85, Application US/09852659A

Patent No. US20020077287A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 28 Human Secreted Proteins

FILE REFERENCE: P2003P4

CURRENT APPLICATION NUMBER: US/09/852,659A

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: 60/265,583

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/152,060

PRIOR FILING DATE: 1998-09-11

PRIOR APPLICATION NUMBER: PCT/US98/04858

PRIOR FILING DATE: 1998-03-12

PRIOR APPLICATION NUMBER: 60/040,762

PRIOR FILING DATE: 1997-03-14

PRIOR APPLICATION NUMBER: 60/040,710

PRIOR FILING DATE: 1997-03-14

PRIOR APPLICATION NUMBER: 60/050,934

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,100

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,357

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/048,189

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: 60/057,765

PRIOR FILING DATE: 1997-09-05

PRIOR APPLICATION NUMBER: 60/048,970

PRIOR FILING DATE: 1997-06-06

PRIOR APPLICATION NUMBER: 60/068,368

PRIOR FILING DATE: 1997-12-19

NUMBER OF SEQ ID NOS: 121

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 85

LENGTH: 121

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (67)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (89)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-852-659A-85

Query Match 60.0%; Score 30; DB 10; Length 121;

Best Local Similarity 85.7%; Pred. No. 82;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

Db 28 EEVVPXG 34

RESULT 21

US-09-992-598-359

Sequence 359, Application US/09992598

Patent No. US20020160384A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavrin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
DB 28 EEVPPGG 34

RESULT 22

US-09-989-293A-359
Sequence 359, Application US/09989293A
Patent No. US20020177164A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730FIC66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10

;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 BEVVPXG 7
| | | | |
Db 28 BEVVEGG 34

RESULT 23

US-09-989-735-359
; Sequence 359, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P27301PC61
CURRENT APPLICATION NUMBER: US/09/989,735
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;

Best Local Similarity 85.7%; Pred. No. 93;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEEVPGX 7

Db 28 EEEVPGG 34

RESULT 24

US-09-990-444-359

; Sequence 359, Application US/09990444

; Publication No. US20020193300A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C19

; CURRENT APPLICATION NUMBER: US/09/990,444

; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
 |||||
Db 28 EEVVPXG 34

RESULT 25

US-09-989-730-359
; Sequence 359, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC69
; CURRENT APPLICATION NUMBER: US/09/989,730
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12

1	60/089532	PRIOR FILING DATE: 1998-06-16
2	60/089533	PRIOR APPLICATION NUMBER: 60/089532
3	60/089534	PRIOR FILING DATE: 1998-06-17
4	60/089535	PRIOR FILING DATE: 1998-06-17
5	60/089536	PRIOR APPLICATION NUMBER: 60/089535
6	60/089537	PRIOR FILING DATE: 1998-06-17
7	60/089538	PRIOR FILING DATE: 1998-06-17
8	60/089539	PRIOR APPLICATION NUMBER: 60/089538
9	60/089540	PRIOR FILING DATE: 1998-06-17
10	60/089541	PRIOR FILING DATE: 1998-06-17
11	60/089542	PRIOR APPLICATION NUMBER: 60/089541
12	60/089543	PRIOR FILING DATE: 1998-06-18
13	60/089544	PRIOR FILING DATE: 1998-06-18
14	60/089545	PRIOR APPLICATION NUMBER: 60/089544
15	60/089546	PRIOR FILING DATE: 1998-06-18
16	60/089547	PRIOR FILING DATE: 1998-06-18
17	60/089548	PRIOR APPLICATION NUMBER: 60/089547
18	60/089549	PRIOR FILING DATE: 1998-06-18
19	60/089550	PRIOR FILING DATE: 1998-06-18
20	60/089551	PRIOR APPLICATION NUMBER: 60/089550
21	60/089552	PRIOR FILING DATE: 1998-06-18
22	60/089553	PRIOR FILING DATE: 1998-06-18
23	60/089554	PRIOR APPLICATION NUMBER: 60/089553
24	60/089555	PRIOR FILING DATE: 1998-06-18
25	60/089556	PRIOR FILING DATE: 1998-06-18
26	60/089557	PRIOR APPLICATION NUMBER: 60/089556
27	60/089558	PRIOR FILING DATE: 1998-06-18
28	60/089559	PRIOR FILING DATE: 1998-06-18
29	60/089560	PRIOR APPLICATION NUMBER: 60/089559
30	60/089561	PRIOR FILING DATE: 1998-06-18
31	60/089562	PRIOR FILING DATE: 1998-06-18
32	60/089563	PRIOR APPLICATION NUMBER: 60/089562
33	60/089564	PRIOR FILING DATE: 1998-06-17
34	60/089565	PRIOR FILING DATE: 1998-06-17
35	60/089566	PRIOR APPLICATION NUMBER: 60/089565
36	60/089567	PRIOR FILING DATE: 1998-06-18
37	60/089568	PRIOR FILING DATE: 1998-06-18
38	60/089569	PRIOR APPLICATION NUMBER: 60/089568
39	60/089570	PRIOR FILING DATE: 1998-06-18
40	60/089571	PRIOR FILING DATE: 1998-06-18
41	60/089572	PRIOR APPLICATION NUMBER: 60/089571
42	60/089573	PRIOR FILING DATE: 1998-06-22
43	60/089574	PRIOR FILING DATE: 1998-06-22
44	60/089575	PRIOR APPLICATION NUMBER: 60/089574
45	60/089576	PRIOR FILING DATE: 1998-06-22
46	60/089577	PRIOR FILING DATE: 1998-06-22
47	60/089578	PRIOR APPLICATION NUMBER: 60/089577
48	60/089579	PRIOR FILING DATE: 1998-06-22
49	60/089580	PRIOR FILING DATE: 1998-06-22
50	60/089581	PRIOR APPLICATION NUMBER: 60/089580
51	60/089582	PRIOR FILING DATE: 1998-06-26
52	60/089583	PRIOR FILING DATE: 1998-06-26
53	60/089584	PRIOR APPLICATION NUMBER: 60/089583
54	60/089585	PRIOR FILING DATE: 1998-06-26
55	60/089586	PRIOR FILING DATE: 1998-06-26
56	60/089587	PRIOR APPLICATION NUMBER: 60/089586
57	60/089588	PRIOR FILING DATE: 1998-07-01
58	60/089589	PRIOR FILING DATE: 1998-07-01
59	60/089590	PRIOR APPLICATION NUMBER: 60/089589
60	60/089591	PRIOR FILING DATE: 1998-06-26
61	60/089592	PRIOR FILING DATE: 1998-06-26
62	60/089593	PRIOR APPLICATION NUMBER: 60/089592
63	60/089594	PRIOR FILING DATE: 1998-06-26
64	60/089595	PRIOR FILING DATE: 1998-06-26
65	60/089596	PRIOR APPLICATION NUMBER: 60/089595
66	60/089597	PRIOR FILING DATE: 1998-06-26
67	60/089598	PRIOR FILING DATE: 1998-06-26
68	60/089599	PRIOR APPLICATION NUMBER: 60/089598
69	60/089600	PRIOR FILING DATE: 1998-06-26
70	60/089601	PRIOR FILING DATE: 1998-06-26
71	60/089602	PRIOR APPLICATION NUMBER: 60/089601
72	60/089603	PRIOR FILING DATE: 1998-06-26
73	60/089604	PRIOR FILING DATE: 1998-06-26
74	60/089605	PRIOR APPLICATION NUMBER: 60/089604
75	60/089606	PRIOR FILING DATE: 1998-06-26
76	60/089607	PRIOR FILING DATE: 1998-06-26
77	60/089608	PRIOR APPLICATION NUMBER: 60/089607
78	60/089609	PRIOR FILING DATE: 1998-06-26
79	60/089610	PRIOR FILING DATE: 1998-06-26
80	60/089611	PRIOR APPLICATION NUMBER: 60/089610
81	60/089612	PRIOR FILING DATE: 1998-06-26
82	60/089613	PRIOR FILING DATE: 1998-06-26
83	60/089614	PRIOR APPLICATION NUMBER: 60/089613
84	60/089615	PRIOR FILING DATE: 1998-06-25
85	60/089616	PRIOR FILING DATE: 1998-06-25
86	60/089617	PRIOR APPLICATION NUMBER: 60/089616
87	60/089618	PRIOR FILING

; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. NO. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

Search completed: June 10, 2003, 14:35:46
Job time : 16.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds

(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-45

Perfect score: 50

Sequence: 1 EEVVPXGXDYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	72.0	363	2 D69551	conserved hypoteth
2	35	70.0	290	2 D98182	6-O-methylguanine-D
3	35	70.0	290	2 AG3104	succinate dehydrog
4	35	70.0	587	2 F81138	disease resistance
5	35	70.0	906	2 T48898	disease resistance
6	35	70.0	908	2 T48899	disease resistance
7	34	68.0	102	2 A42452	V1 protein - tobac
8	34	68.0	156	2 S54619	hypothetical prote
9	34	68.0	247	2 A96001	conserved hypoteth
10	34	68.0	257	2 A96546	unknown protein [i
11	34	68.0	394	2 F82491	ferrisiderophore r
12	34	68.0	433	2 H87660	peptidoglycan-bind
13	34	68.0	2747	2 B49132	fat facets (faf) s
14	33	66.0	124	1 VKLJ51	trans-regulatory s
15	33	66.0	165	2 AG1272	thiol peroxidases
16	33	66.0	165	2 AH1635	thiol peroxidases
17	33	66.0	196	2 AD0454	conserved hypoteth
18	33	66.0	225	2 S57810	hypothetical prote
19	33	66.0	327	2 S40753	hypothetical prote
20	33	66.0	421	1 DERTCM	acyl-coA dehydroge
21	33	66.0	440	2 H72784	probable alkaline
22	33	66.0	1028	2 AF3286	ATP-dependent DNA
23	33	66.0	1088	2 D82246	probable chitinase
24	33	66.0	1150	2 T20173	hypothetical prote
25	32	64.0	99	2 S00210	plastocyanin b - L
26	32	64.0	155	2 S38255	plastocyanin b pre
27	32	64.0	168	2 S58208	conserved hypoteth
28	32	64.0	196	2 AI0931	hypothetical prote
29	32	64.0	301	2 F89957	hypothetical prote

ALIGNMENTS

RESULT 1

D69551 conserved hypothetical protein AF2411 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: D69551

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.;

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaach, P.; Kalne, B.P.; Syk

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing a

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69551

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-363 <KLE>

A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AB91255.1; PID:g

Query Match 72.0%; Score 36; DB 2; Length 363;

Best Local Similarity 54.5%; Pred. No. 9;

30	64.0	307	2 F84330
31	64.0	314	2 AH1912
32	64.0	357	1 G69290
33	64.0	366	2 G89350
34	64.0	425	2 T24111
35	64.0	427	2 F64064
36	64.0	565	2 E86665
37	64.0	632	2 H84350
38	64.0	672	2 G88651
39	64.0	1474	2 F69009
40	64.0	1472	2 T31308
41	64.0	6658	2 T13931
42	62.0	117	2 A69487
43	62.0	203	2 H97247
44	62.0	233	2 E72330
45	62.0	296	2 F72745
46	62.0	319	2 S03833
47	62.0	395	2 H84113
48	62.0	421	1 DEHUCM
49	62.0	496	2 S76296
50	62.0	587	2 D81881
51	62.0	622	2 S15009
52	62.0	630	2 F85074
53	62.0	653	2 D82352
54	62.0	840	2 AG0526
55	62.0	846	2 S57580
56	62.0	1741	2 S74910
57	62.0	13055	2 T16580
58	61.0	668	2 JQ2356
59	60.0	21	2 B49042
60	60.0	97	2 JW0011
61	60.0	97	2 A99427
62	60.0	128	2 A90471
63	60.0	165	2 D89493
64	60.0	175	2 S36749
65	60.0	180	2 AG0504
66	60.0	184	2 B86192
67	60.0	184	2 EG0335
68	60.0	204	2 S66082
69	60.0	231	1 ISECP4
70	60.0	231	2 A90637
71	60.0	231	2 A85488
72	60.0	232	2 AB0515
73	60.0	232	2 C71908
74	60.0	232	2 C64606
75	60.0	248	1 ISEB4T

hypothetical prote
hypothetical prote
probable hexosyltr
L-lactate dehydrog
hypothetical prote
toIB protein - Hae
ABC transporter AT
oligopeptidase [im
protein B0212.3 [i
probable membrane
hypothetical 367K
projectin - fruit
response regulator
S-adenosylmethioni
conserved hypoteth
hypothetical prote
hypothetical prote
hypothetical prote
acyl-CoA dehydroge
hypothetical prote
probable succinate
hypothetical prote
hypothetical prote
iron(III) ABC tran
penicillin-binding
penicillin-binding
hemolysin - Synecch
hypothetical prote
capsid protein - f
Ig heavy chain V r
plastocyanin - car
partial transposas
hypothetical prote
hypothetical prote
transcription fact
fimbrial chain [im
hypothetical prote
hypothetical prote
hypothetical prote
general stress pro
L-ribulose-phospha
L-ribulose-5-phosp
L-ribulose-5-phosp
3-oxoacid CoA-tran
3-oxoadipate CoA-t
L-ribulose-phospha

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXS 11
| : | | | : | |
Db 120 ENIVPGIDFS 130

RESULT 2
D98182
O6-methylguanine-DNA methyltransferase PA2118 [imported] - Agrobacterium tumefaciens (strain C)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98182
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens 294, 2323-2328, 2001
A:Reference number: A97359; PMID:11743194
A:Accession: D98182
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK88982.1; PID:g15158766; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_818
A:Map position: linear chromosome

Query Match 70.0%; Score 35; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXD 10
| : | | | : | |
Db 9 EDITPIGSDY 18

RESULT 3
AG3104
6-O-methylguanine-DNA methyltransferase [imported] - Agrobacterium tumefaciens (strain C)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG3104
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG3104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: GB:AE008689; PIDN:AA45253.1; PID:g17742937; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: ada
A:Map position: linear chromosome

Query Match 70.0%; Score 35; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXD 10
| : | | | : | |
Db 9 EDITPIGSDY 18

RESULT 4
F81138
succinate dehydrogenase, flavoprotein chain NMB0950 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81138
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eise
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty,
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli,
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC5
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <PET>
A:Cross-references: GB:AE002446; GB:AE002098; NID:g7226185; PIDN:AAF41356.1; PID:g
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0950
C:Superfamily: fumarate reductase flavoprotein; 3-oxosteroid 1-dehydrogenase homol

Query Match 70.0%; Score 35; DB 2; Length 587;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXD 10
| : | | | : | |
Db 366 EVVVPQGEDY 375

RESULT 5
T48898
disease resistance protein RPP8 [validated] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 21-Jul-2000
C:Accession: T48898
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.;
Plant Cell 10, 1861-1874, 1998
A:Title: Intragenic recombination and diversifying selection contribute to the evol
A:Reference number: Z42999; MUID:99030193; PMID:9811794
A:Accession: T48898
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-906 <MCD>
A:Cross-references: EMBL:AF089710; NID:g3928861; PIDN:AAC83165.1; PID:g3928862
A:Experimental source: Landsberg erecta
C:Genetics:
A:Gene: RPP8
A:Introns: 293/1; 342/1
C:Function:
A:Description: promotes resistance to Peronospora parasitica

Query Match 70.0%; Score 35; DB 2; Length 906;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXD 10
| : | | | : | |
Db 881 EKLVPQGEDY 890

RESULT 6
T48899
disease resistance protein rpp8 [similarity] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 21-Jul-2000
C:Accession: T48899
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.;
Plant Cell 10, 1861-1874, 1998
A:Title: Intragenic recombination and diversifying selection contribute to the evol
A:Reference number: Z24999; MUID:99030193; PMID:9811794
A:Accession: T48899
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-908 <MCD>

A:Cross-references: EMBL:AF089711; NID:g3901293; PIDN:AAC78631.1; PID:g3901294
 A:Experimental source: Columbia
 C:Genetics:

A:Gene: rpp8
 A:Introns: 293/1; 342/1
 C:Function:

A:Description: susceptible allele of a gene that promotes resistance to Peronospora para

Query Match 70.0%; Score 35; DB 2; Length 908;
 Best Local Similarity 60.0%; Pred. No. 40;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 : : : : :
 Db 883 EKLVPAGEDY 892

RESULT 7

A42452

V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virolology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 68.0%; Score 34; DB 2; Length 102;
 Best Local Similarity 60.0%; Pred. No. 6;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXGXDYS 11
 : : : : :
 Db 7 QVVPXGXGINS 16

RESULT 8

S54619

hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae).

N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C:Accession: S54619; S66879

R:de Haan, M.; Maarse, A.C.; Grivell, L.A.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54619

A:Accession: S54619

A:Molecule type: DNA

A:Residues: 1-156 <DEH>

A:Cross-references: EMBL:X87331; NID:gl041652; PIDN:CAA60762.1; PID:g829123

R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66877

A:Accession: S66879

A:Molecule type: DNA

A:Residues: 1-156 <DEW>

A:Cross-references: EMBL:Z74920; NID:gl420109; PIDN:CAA99201.1; PID:gl420111; MIPS:YOR01

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0005539

A:Map position: 13R

C:Superfamily: hypothetical protein YOR013w

Query Match 68.0%; Score 34; DB 2; Length 156;
 Best Local Similarity 66.7%; Pred. No. 9.6;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXGXDYS 10

Db 50 EVMPLGMDY 58
 : : : : :
 : : : : :

RESULT 9

A96001

conserved hypothetical protein, homolog to osmotically inducible sensory protein SM

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Barloy-Hu

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: A96001

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49673.1; PID:gl5141160; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hu

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher,

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Le

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Y

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMB21444

A:Genome: plasmid

Query Match 68.0%; Score 34; DB 2; Length 247;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 : : : : :
 Db 48 EDVEPRGADY 57

RESULT 10

A96546

unknown protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A96546

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewa

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mar

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Ta

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: A96546

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-257 <STO>

A:Cross-references: GB:AE005173; NID:gl1094688; PIDN:AAG29624.1; GSPDB:GN00141

C:Genetics:

A:Gene: F8A12.12

A:Map position: 1

Query Match 68.0%; Score 34; DB 2; Length 257;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXGXDYS 11

```

Db      217 EELKAGADYS 227
||||: | |||

RESULT 11
F82491
ferrisiderophore reductase VCA0183 [imported] - Vibrio cholerae (strain N16961 serogroup
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000. #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82491
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82491
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <HEI>
A:Cross-references: GB:AE004358; GB:AE003853; NID:g9657566; PIDN:AAF96096.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0183
A:Map position: 2
C:Superfamily: flavohemoglobin; cytochrome-b5 reductase homology; globin homology

Query Match      68.0%; Score 34; DB 2; Length 394;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 EVVPXGXDY 10
|| || || ||
Db      194 EVTPEGSDY 202

RESULT 12
H87660
peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87660
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87660
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <STO>
A:Cross-references: GB:AE005673; NID:g13425020; PIDN:AAK25284.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3322

Query Match      68.0%; Score 34; DB 2; Length 433;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 EVVPXGXDYS 11
| : | | | |
Db      266 EVILPGFDYS 276

RESULT 13
B49132
fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
C:Accession: B49132; A49132
R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
Development 116, 985-1000, 1992

```

```

A:Title: The fat facets gene is required for Drosophila eye and embryo development
A:Reference number: A49132; MUID:93202020; PMID:1295747
A:Contents: isogenic st
A:Accession: B49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2747 <FIS>
A:Cross-references: GB:I04959; NID:gl57411; PIDN:AAF01345.1; PID:g6013474
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
A:Accession: A49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2704 'VT', 2707 'ANNV' <FI2>
A:Cross-references: GB:I04958; NID:gl57410; PIDN:AAF01346.1; PID:g6013475
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
C:Keywords: alternative splicing

Query Match      68.0%; Score 34; DB 2; Length 2747;
Best Local Similarity 54.5%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGXDXS 11
| : | | | |
Db      1394 EVIVPDGQDFS 1404

RESULT 14
VKLSJI
trans-regulatory splicing protein - simian immunodeficiency virus SIVcpz
N:Alternate names: anti-repression trans-activator; art protein; rev protein; trs
C:Species: simian immunodeficiency virus SIVcpz
A>Note: host Pan troglodytes (chimpanzee)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09988
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09988
A>Status: nucleic acid sequence, not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-124 <HUE>
A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36405.1; PID:g763085
C:Genetics:
A:Gene: rev; trs; art
A:Introns: 27/1
A:Superfamily: AIDS trans-regulatory splicing protein
C:Keywords: AIDS; immunodeficiency; splicing protein; transcription regulation

Query Match      66.0%; Score 33; DB 1; Length 124;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 EVVPXGXDXS 11
| : | | | |
Db      107 ETVPAGGNYS 116

RESULT 15
AG1272
thiol peroxidases homolog lmo1583 [imported] - Listeria monocytogenes (strain EGD-e
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C:Accession: AG1272
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Blo
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsih
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Maqueno, E.; Maitournam, A
ok, C.; Schluter, T.; Smeos, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weh
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

```

A:Accession: AG1272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99661.1; PID:g16411012; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lml583
C:Superfamily: thioredoxin peroxidase

Query Match 66.0%; Score 33; DB 2; Length 165;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10
|||||
Db 144 EVVPEGSDH 152

RESULT 16
AH1635
thiol peroxidases homolog lin1625 [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C:Accession: AH1635
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1635
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96856.1; PID:g16414112; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1625
C:Superfamily: thioredoxin peroxidase

Query Match 66.0%; Score 33; DB 2; Length 165;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10
|||||
Db 144 EVVPEGSDH 152

RESULT 17
AD0454
conserved hypothetical protein YPO3732 [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AD0454
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0454
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93200.1; PID:g15981648; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3732

Query Match 66.0%; Score 33; DB 2; Length 196;

Best Local Similarity 45.5%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
|||||
Db 76 EEAIPSSDDYA 86

RESULT 18
S57810
hypothetical protein precursor (clone TP11) - tomato
C:Species: *Lycopersicon esculentum* (tomato)
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S57810
R:Milligan, S.B.; Gasser, C.S.
Plant Mol. Biol. 28, 691-711, 1995
A:Title: Nature and regulation of pistil-expressed genes in tomato.
A:Reference number: S57808; MUID:95375233; PMID:7647301
A:Accession: S57810
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-225 <MIL>
A:Cross-references: EMBL:U20592; NID:g924625; PIDN:AAA80497.1; PID:g924626
C:Superfamily: plant Kunitz-type proteinase inhibitor

Query Match 66.0%; Score 33; DB 2; Length 225;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
|||||
Db 32 DEVVPNGKTYA 42

RESULT 19
S40753
hypothetical protein C15H7.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C:Accession: S40753
R:Smith, A.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40750
A:Accession: S40753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <SMI>
A:Cross-references: EMBL:222173; NID:g297944; PID:g297948
C:Genetics:
A:Introns: 14/3; 59/2; 115/2; 188/2; 238/3

Query Match 66.0%; Score 33; DB 2; Length 327;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
|||||
Db 175 KEVVPNGGDKS 185

RESULT 20
DERTCM
acyl-CoA dehydrogenase (EC 1.3.99.3) precursor, medium-chain-specific, mitochondrial
N:Alternate names: acyl dehydrogenase, medium-chain-specific
C:Species: *Rattus norvegicus* (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 01-Dec-2000
C:Accession: A28436; S15128
R:Matsubara, Y.; Kraus, J.P.; Ozasa, H.; Glassberg, R.; Finocchiaro, G.; Ikeda, Y.
J. Biol. Chem. 262, 10104-10108, 1987
A:Title: Molecular cloning and nucleotide sequence of cDNA encoding the entire pre
A:Reference number: A28436; MUID:87280028; PMID:3611054
A:Accession: A28436
A:Molecule type: mRNA

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDXS 11

I: | | | |

Db 562 VLPVGIDYS 570

RESULT 25

S00210

plastocyanin b - Lombardy poplar

C:Species: Populus nigra var. italica (Lombardy poplar)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 20-Apr-2000

C:Accession: S00210

R:Dimitrov, M.I.; Egorov, C.A.; Donchev, A.A.; Atanasov, B.P.

FEBS Lett. 226, 17-22, 1987

A:Title: Complete amino acid sequence of poplar plastocyanin b.

A:Reference number: S00210

A:Accession: S00210

A:Molecule type: protein

A:Residues: 1-99 <DIM>

C:Superfamily: plastocyanin

C:Keywords: chloroplast; copper; electron transfer; metalloprotein

F:37,84,87,92/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match

64.0%; Score 32; DB 2; Length 99;

Best Local Similarity 54.5%; Pred. NO. 16;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDXS 11

I: | | | |

Db 43 EDVPSGVDVS 53

Search completed: June 10, 2003, 13:49:18

Job time : 12.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-45
Perfect score: 50
Sequence: 1 EEVVPXGXDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	35	70.0	1058	1	CARB_FUSNN
2	34	68.0	102	1	Y1LK_TYDVA
3	34	68.0	394	1	HMPA_VIBCH
4	34	68.0	2747	1	EAF_DROME
5	33	66.0	124	1	REV_SIVCZ
6	33	66.0	327	1	YK14_CAEEL
7	33	66.0	421	1	ACDM_RAT
8	32	64.0	154	1	PLAS_ORYSA
9	32	64.0	155	1	PLAS_HORVU
10	32	64.0	168	1	PLAT_POPNI
11	32	64.0	196	1	YUNG_SALTY
12	32	64.0	427	1	TOUB_HAEIN
13	32	64.0	1499	1	A10C_HUMAN
14	32	64.0	3174	1	CHAC_HUMAN
15	31	62.0	233	1	HIS9_THEMEA
16	31	62.0	319	1	YHAI_CRYPA
17	31	62.0	421	1	ACDM_HUMAN
18	31	62.0	421	1	ACDM_PIG
19	31	62.0	562	1	TR2M_ERWHE
20	30	60.0	97	1	PLAS_DAUCA
21	30	60.0	121	1	TKNK_HUMAN
22	30	60.0	175	1	HES3_RAT
23	30	60.0	203	1	CTC_BACSU
24	30	60.0	231	1	ARAD_ECOLI
25	30	60.0	231	1	ARAD_SALTY
26	30	60.0	232	1	SCOA_HELJP
27	30	60.0	232	1	SCOA_HELJP
28	30	60.0	259	1	OVUH_LYMSY
29	30	60.0	421	1	AMP2_YEAST
30	30	60.0	421	1	ECB2_HALEL
31	30	60.0	423	1	ECB1_HALEL
32	30	60.0	430	1	FOLC_BACSU
33	30	60.0	457	1	Z185_HUMAN

34	30	60.0	469	1	LET1_KLULA
35	30	60.0	478	1	AMP2_HUMAN
36	30	60.0	478	1	AMP2_MOUSE
37	30	60.0	478	1	AMP2_RAT
38	30	60.0	478	1	AMP2_MOUSE
39	30	60.0	530	1	GSR2_HUMAN
40	30	60.0	530	1	APHF_XANCH
41	30	60.0	556	1	APLY_APLKU
42	30	60.0	583	1	HEMO_BRARE
43	30	60.0	890	1	ECNS_CLOPE
44	30	60.0	1049	1	SPS_ORYSA
45	30	60.0	1068	1	SPS2_CRAPL
46	30	60.0	1081	1	SPS2_CRAPL
47	30	60.0	1176	1	NIR_NEUCR
48	30	60.0	1401	1	RPOC_VIBCH
49	30	60.0	1501	1	SNQ2_YEAST
50	30	60.0	1729	1	TABP_HUMAN
51	30	60.0	2717	1	ZEPI_HUMAN
52	29.5	59.0	472	1	POLG_PEMVM
53	29	58.0	98	1	ET2A_XENLA
54	29	58.0	98	1	PLAS_ENTPR
55	29	58.0	98	1	PLAS_ULVAR
56	29	58.0	99	1	PLAS_RUMOB
57	29	58.0	99	1	PLAS_TOBAC
58	29	58.0	113	1	Y011_BPT4
59	29	58.0	175	1	HES3_MOUSE
60	29	58.0	230	1	YCAP_ECOLI
61	29	58.0	242	1	PSB4_XENLA
62	29	58.0	258	1	YK10_MOUSE
63	29	58.0	261	1	YK10_MOUSE
64	29	58.0	338	1	MTBA_METBA
65	29	58.0	354	1	VGLI_VZVD
66	29	58.0	388	1	FBM3_CAEEL
67	29	58.0	389	1	SAT_AERPE
68	29	58.0	412	1	DOK2_HUMAN
69	29	58.0	421	1	ACDM_MOUSE
70	29	58.0	424	1	MS11_ARATH
71	29	58.0	435	1	HEM1_DESYM
72	29	58.0	435	1	HEM1_DESYM
73	29	58.0	471	1	TMLH_NEUCR
74	29	58.0	472	1	ET2B_XENLA
75	29	58.0	477	1	CAP2_RAT

ALIGNMENTS

RESULT 1	CARB_FUSNN	STANDARD	PRT: 1058 AA.
AC	Q8RG86:		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Carbamoyl-phosphatase synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).		
DE	Carb OR FN0422.		
GN	Fusobacterium nucleatum (subsp. nucleatum).		
OC	Bacteria; Fusobacteria; Fusobacterium.		
OX	NCBI_TaxID=76856;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 25586;		
RC	MEDLINE=21886394; PubMed=11889109;		
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyripides N., Overbeek R.;		
RA	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586."		
RT	Q95877 halomonas e		
RL	Q05865 bacillus su		
CC	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP +		

phosphate + L-glutamate + carbamoyl phosphate.
 -!- COFACTOR: Binds three manganese ions (by similarity).
 -!- PATHWAY: Arginine biosynthesis.
 -!- SUBUNIT: Pyrimidine biosynthesis; first step.
 promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (by similarity).
 -!- SIMILARITY: BELONGS TO THE CARB FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE010554; AAL94625.1; ALT_INIT.
 InterPro; IPR005483; CPase_L.
 InterPro; IPR005479; CPase_L_D2.
 InterPro; IPR005480; CPase_L_D3.
 InterPro; IPR005481; CPase_L_N.
 InterPro; IPR004362; MGS-like.
 Pfam; PF00289; CPase_L_chain; 2.
 Pfam; PF02786; CPase_L_D2; 2.
 Pfam; PF02787; CPase_L_D3; 1.
 Pfam; PF02142; MGS; 1.
 PRINTS; PR00098; CPASE.
 PROSITE; PS00866; CPASE_1; 2.
 PROSITE; PS00867; CPASE_2; 2.
 Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 ATP-binding; Manganese; Complete proteome.
 FT DOMAIN 1 401
 FT DOMAIN 402 546
 FT DOMAIN 547 929
 FT DOMAIN 930 1058
 FT REPEAT 1 546
 FT REPEAT 547 1058
 FT NP_BIND 153 210
 FT NP_BIND 302 352
 FT METAL 284 284
 FT METAL 298 298
 FT METAL 300 300
 FT METAL 820 820
 FT METAL 832 832
 SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;

Query Match 70.0%; Score 35; DB 1; Length 1058;
 Best Local Similarity 60.0%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDYS 11
 Db 190 EIVPGLNYS 199

RESULT 2
 Y11K_TYDVA
 ID Y11K_TYDVA STANDARD; PRT; 102 AA.
 AC F31619;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE Hypothetical 11.2 kDa protein.
 GN Y1.
 OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OX NCBI_TaxID=31599;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92188538; PubMed=1546458;
 RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses
 RT infecting monocotyledonous plants.";
 RL Virology 187:633-642(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81103; AAA47947.1; -.
 DR PIR; A42452; A42452.
 DR InterPro; IPR002621; Gemini_mov.
 DR Pfam; PF01708; Gemini_mov; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF5B67 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 102;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDYS 11
 Db 7 QVVPNGINYS 16

RESULT 3
 HMPA_VIBCH
 ID HMPA_VIBCH STANDARD; PRT; 394 AA.
 AC Q9KMY3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
 GN HMP OR VCA0183.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";
 RT Nature 406:477-483(2000).
 RL Nature 406:477-483(2000).
 CC -!- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS; ONE IS A HEME-CONTAINING OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
 CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN FLAVOHEMOPROTEIN SUBFAMILY.
 CC -!- SIMILARITY: TO A NUMBER OF FAD/NAD(P) FLAVOPROTEIN OXIDOREDUCTASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE004358; AAF96096.1; -.
 DR HSSP; P39662; 1CQX.

DR TIGR; VCA0183; -
 DR InterPro; IPR001834; Cyt_B5_reductase.
 DR InterPro; IPR001709; FPN_cyt_redctse.
 DR InterPro; IPR000971; Globin.
 DR InterPro; IPR001433; Oxred_FAD/NAD(P).
 DR InterPro; IPR001221; Phe_hydroxylase.
 DR Pfam; PF00042; globin; 1.
 DR Pfam; PF00175; NAD_binding; 1.
 DR Pfam; PF00970; FAD_binding_6; 1.
 DR PRINTS; PR00371; FPNCR.
 DR PRINTS; PR00410; PHEHYDRLASE.
 DR PROSITE; PS01033; GLOBIN; 1.
 KW Oxidoreductase; NADP; Heme; Flavoprotein; FAD; Iron transport;
 Oxygen transport; Transpore; Complete proteome.
 KW DOMAIN 1 136
 FT METAL 53 53
 FT FT
 FT METAL 85 85
 FT FT
 FT NP_BIND 268 273 NADP (RIBOSE PART) (BY SIMILARITY).
 FT SEQUENCE 394 AA; 44191 MW; 4490FAE28823A CRC64;
 Query Match 68.0%; Score 34; DB 1; Length 394;
 Best Local Similarity 66.7%; Pred. No. 9.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 EVVPXGXDY 10
 ||| |||
 Db 194 EVTPGSDY 202
 RESULT 4
 FAF_DROME
 ID FAF_DROME STANDARD; PRT; 2747 AA.
 AC P55824;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
 DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease
 DE FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
 GN FAF.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93202020; PubMed=1295747;
 RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;
 RT The fat facets gene is required for Drosophila eye and embryo
 development.";
 RL Development 116:985-1000(1992).
 CC -!- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A
 CC ROLE IN COMPOUND EYE ASSEMBLY AND OOGENESIS RESPECTIVELY. IN THE
 CC LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS
 CC PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY
 CC CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR
 CC NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND
 CC COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR
 CC FUNCTION.
 CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L04959; AAF01345.1; -;
 CC EMBL; L04958; AAF01346.1; -;
 CC MEROPS; C19.007; -;
 DR FlyBase; FBgn0005632; faf.
 DR InterPro; IPR001394; UCH-2.
 DR Pfam; PF00442; UCH-1; 1.
 DR Pfam; PF00443; UCH-2; 1.
 DR PROSITE; PS00972; UCH_2_1; 1.
 DR PROSITE; PS00973; UCH_2_2; 1.
 DR PROSITE; PS02035; UCH_2_3; 1.
 KW Ubl conjugation pathway; Hydrolase; Thiol protease;
 Developmental protein; Vision; Alternative splicing.
 KW ACT_SITE 1677 1677 BY SIMILARITY.
 FT ACT_SITE 1978 1978 BY SIMILARITY.
 FT ACT_SITE 1986 1986 BY SIMILARITY.
 FT VARSPLIC 2705 2747 KCRVVIKKLVESKDEEDATSATTAATTEVTTSPATAS
 FT VARIANT 2725 2725 ORQQL -> VTRANNV (IN SHORT ISOFORM).
 FT SEQUENCE 2747 AA; 307954 MW; 1D97659F7A7B2ADE CRC64;
 Query Match 68.0%; Score 34; DB 1; Length 2747;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGXNDYS 11
 ||| |||
 Db 1394 EVIPVDGQDFS 1404
 RESULT 5
 REV_SIVCZ
 ID REV_SIVCZ STANDARD; PRT; 124 AA.
 AC P17280;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE REV protein (Anti-repression transactivator protein) (ART/TRS).
 DE REV.
 GN REV.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90259077; PubMed=2188136;
 RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
 RT "Genetic organization of a chimpanzee lentivirus related to HIV-1.";
 RL Nature 345:356-359(1990).
 CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
 CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
 CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
 CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
 CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X52154; CAA36405.1; -;
 CC PIR; S09988; VKLJSI.
 DR HIV; X52154; REVSCPZ.
 DR InterPro; IPR000625; REV_protein.
 DR Pfam; PF00424; REV; 1.
 KW Transcription regulation; AIDS; phosphorylation; Nuclear protein.
 SQ SEQUENCE 124 AA; 13701 MW; F5877D1BDF65A7B2 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 124;
 Best Local Similarity 60.0%; Pred. No. 4.6;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGXDXS 11
 |||||
 Db 107 ETVPAGNGYS 116

RESULT 6

YK14_CAEEL STANDARD; PRT; 327 AA.

AC P34338;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C15H7.4 in chromosome III.
 GN C15H7.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN-Bristol N2;

CC MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wooldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z22173; CAA80126.1; -

DR PIR; S40753; S40753.

DR WormPep; C15H7.4; CE00082.

KW Hypothetical protein.

SQ SEQUENCE 327 AA; 35566 MW; 716BC2BDD2E9607E CRC64;

Query Match 66.0%; Score 33; DB 1; Length 327;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGXDXS 11
 |||||
 Db 175 KEVPPNGGDKS 185

RESULT 7

ACDM_RAT STANDARD; PRT; 421 AA.

AC P08503;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
 (EC 1.3.99.3) (MCAD).

DE

GN ACADM.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=87280028; PubMed=3611054;

RA Matsubara Y., Kraus J.P., Ozasa H., Glassberg R., Finocchiaro G.,

RA Ikeda Y., Mole J., Rosenberg L.E., Tanaka K.;

RT "Molecular cloning and nucleotide sequence of cDNA encoding the

RT entire precursor of rat liver medium chain acyl coenzyme A

RT dehydrogenase";

RL J. Biol. Chem. 262:10104-10108(1987).

CC -!- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO

CC 16.

CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced

CC ETF.

CC -!- COFACTOR: FAD.

CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first

CC step.

CC -!- SUBUNIT: HOMOTETRAMER.

CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-CoA DEHYDROGENASES

CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN

CC TISSUES.

CC -!- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; J02791; AAA40670.1; -

DR PIR; A28436; DERTCM.

DR HSP; P11310; LEGD.

DR InterPro; IPR001552; Acyl-CoA_dh.

DR Pfam; PF00441; Acyl-CoA_dh; 1.

DR Pfam; PF02770; Acyl-CoA_dh_M; 1.

DR Pfam; PF02771; Acyl-CoA_dh_N; 1.

DR PROSITE; PS00072; ACYL_COA_DH_1; 1.

DR PROSITE; PS00073; ACYL_COA_DH_2; 1.

KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;

KW Mitochondrion; Transit peptide.

FT TRANSIT 1 25 MITOCHONDRION.

FT CHAIN 26 421 ACYL-CoA DEHYDROGENASE, MEDIUM-CHAIN

FT SPECIFIC.

FT ACT_SITE 193 193 FORMS A HYDROGEN-BOND WITH THE FLAVIN

FT N(5) OF THE FAD COFACTOR (BY SIMILARITY).

FT ACT_SITE 401 401 BASE (BY SIMILARITY).

SQ SEQUENCE 421 AA; 46555 MW; 2CF076F8C919BDE8 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 421;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGXDX 10

|||||

Db 58 EEIIPVADPY 67

RESULT 8

PLAS_ORYSA STANDARD; PRT; 154 AA.

ID P20423; Q58B8;

DT 01-FEB-1991 (Rel. 17, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Plastocyanin, chloroplast precursor.

DE

PETE.
 Oryza sativa (Rice).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=cv. Ilpoom; TISSUE=Leaf;
 Lee J.-S.;
 "Molecular cloning and characterization of plastocyanin precursor in
 rice";
 Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 SEQUENCE OF 58-154.
 STRAIN=cv. Japonica;
 MEDLINE=89386623; PubMed=2780537;
 Yano H., Kamo M., Tsugita A., Aso K., Nozu Y.;
 "The amino acid sequence of plastocyanin from rice (*Oryza sativa*,
 subspecies japonica).";
 RT Protein Seq. Data Anal. 2:385-389(1989).
 CC -I- FUNCTION: Participates in electron transfer between P700 and the
 cytochrome b6-f complex in photosystem I.
 CC -I- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -I- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; AF093636; AAC78108.1; -
 PIR; S06105; S06105.
 PIR; J0352; J0352.
 HSP; P00289; 2PCF.
 InterPro; IPR001235; BlueCu_1.
 InterPro; IPR001235; Copper_1blue.
 Pfam; PF00127; copper-bind; 1.
 PRIN5; PR00156; COPPERBLU1E.
 PRODOM; PD001235; Copper_blue; 1.
 PROSITE; PS00196; COPPER_BLUE; 1.
 Chroloplast; Electron transport; Copper; Thylakoid; Membrane;
 Transit peptide.
 TRANSIT 1 57
 CHLOROPLAST.
 FT CHAIN 58 154
 PLASTOCYANIN.
 FT DOMAIN 58 154
 COPPER (BY SIMILARITY).
 FT METAL 94 94
 COPPER (BY SIMILARITY).
 FT METAL 139 139
 COPPER (BY SIMILARITY).
 FT METAL 142 142
 COPPER (BY SIMILARITY).
 FT METAL 147 147
 COPPER (BY SIMILARITY).
 SQ SEQUENCE 154 AA; 15577 MW; E45725D25B5F400D CRC64;

 Query Match 64.0%; Score 32; DB 1; Length 154;
 Best Local Similarity 54.5%; Pred. No. 9.6;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDSYs 11
 I: | | | | |
 Db 100 EDAPVSGVDVS 110

 RESULT 9.
 PLAS_HORVU
 ID PLAS_HORVU STANDARD; PRT; 155 AA.
 AC P08248;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin, chloroplast precursor.

DE Plastocyanin B, chloroplast precursor.
 GN PEST.
 OS Populus nigra (Lombardy poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Italica; TISSUE=Leaf;
 RA Reichert J., Jenzelewski V., Haehnel W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 70-168.
 RC STRAIN=cv. Italica;
 RA Dmitrov M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;
 RT "Complete amino acid sequence of poplar plastocyanin b.";
 RL FEBS Lett. 226:17-22(1987).
 CC -1- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF
 CC POPULAR PLASTOCYANINS A AND B.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z50186; CAA90565.1; -;
 DR PIR; S00210; S00210.
 DR HSP; P00299; 1PLC.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_blue.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLU6.
 DR PRODOM; PD001235; Copper_blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 DR Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide; Multigene family.
 FT TRANSIT 1 69
 FT CHAIN 70 168
 FT DOMAIN 70 168
 FT METAL 106 106
 FT METAL 153 153
 FT METAL 156 156
 FT METAL 161 161
 FT METAL 161 161
 FT COPPER.
 SQ SEQUENCE 168 AA; 16981 MW; F20DAGEA2038AEEA CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 168;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db 112 EDAPVSGVDVS 122
 RESULT 11
 ID YJAG_SALTY STANDARD; PRT; 196 AA.
 AC Q9L912;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yJag.
 GN YJAG OR STM4169 OR STMFI.23 OR STY3716.
 OS Salmonella typhimurium, and

OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE HI0431.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF170176; AAF33519.1; -;
 DR EMBL; AE008894; AAL22997.1; -;
 DR EMBL; AL627279; CAD09475.1; -;
 DR StyGene; SG????; YJag.
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 196 AA; 22704 MW; E40049CE5EE564150 CRC64;
 SQ
 Query Match 64.0%; Score 32; DB 1; Length 196;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 Db 76 EEATPAADDY 85
 RESULT 12
 ID TOLB_HAEIN STANDARD; PRT; 427 AA.
 AC P44677; P94811;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE TolB protein precursor.
 GN TOLB OR HI0382.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RT Science 269:496-512(1995).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=1479;
 RC MEDLINE=97080550; PubMed=8921895;
 RX Sen K., Sikkema D.J., Murphy T.F.;
 RA "Isolation and characterization of the Haemophilus influenzae tolQ,
 RT tolR, tolA and tolB genes.";
 RL Gene 178:75-81(1996).
 CC -!- FUNCTION: INVOLVED IN THE TONB-INDEPENDENT UPTAKE OF PROTEINS
 CC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -!- SIMILARITY: BELONGS TO THE TOLB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32722; AAC22040.1; -;
 DR EMBL; U32470; AAC44597.1; -;
 DR HSSP; P19935; 1CRZ.
 DR TIGR; HT0382; -;
 KW Transport; Protein transport; Periplasmic; Signal; Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 427 TOLB PROTEIN.
 FT VARIANT 6 6 R -> H (IN STRAIN 1479).
 FT VARIANT 14 14 V -> I (IN STRAIN 1479).
 FT VARIANT 17 19 VGS -> ITH (IN STRAIN 1479).
 FT VARIANT 21 21 A -> V (IN STRAIN 1479).
 FT VARIANT 79 79 R -> H (IN STRAIN 1479).
 FT VARIANT 129 129 T -> A (IN STRAIN 1479).
 FT VARIANT 160 160 A -> G (IN STRAIN 1479).
 FT VARIANT 237 237 A -> T (IN STRAIN 1479).
 FT VARIANT 322 322 S -> N (IN STRAIN 1479).
 FT VARIANT 326 326 A -> V (IN STRAIN 1479).
 FT VARIANT 328 328 A -> S (IN STRAIN 1479).
 SQ SEQUENCE 427 AA; 44967 MW; 0882201AEE9254B9 CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 427;
 Best Local Similarity 60.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 EVVPGXGDYS 11
 Db 103 QVVPNGNGYS 112
 RESULT 13
 ALOC_HUMAN STANDARD; PRT; 1499 AA.
 AC 060312; Q96914;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)
 DE (Aminophospholipid translocase VC).
 GN ATP10C OR ATPVC OR KIAA0566.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21225279; PubMed=11326269;
 RA Meguro M., Kashiwagi A., Mitsuwa K., Nakao M., Kondo I., Saitoh S.,
 RA Oshimura M.;
 RA "A novel maternally expressed gene, ATP10C, encodes a putative
 RT aminophospholipid translocase associated with Angelman syndrome.";
 RL Nat. Genet. 28:19-20(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21313119; PubMed=11353404;
 RA Herzog L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;
 RA "The human aminophospholipid-transporting ATPase gene ATP10C maps
 RT adjacent to UBE3A and exhibits similar imprinted expression.";
 RL Am. J. Hum. Genet. 68:1501-1503(2001).
 RN [3]
 RP SEQUENCE OF 337-1499 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:331-39(1998).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + phosphate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Wide expression, with highest levels in
 CC kidney, followed by lung, brain, prostate, testis, ovary, and
 CC small intestine.
 CC -!- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome
 CC (AS), also known as 'happy puppet syndrome'.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
 CC ATPASES). SUBFAMILY IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB051358; BAB47392.1; -;
 DR EMBL; AY029504; AAK33100.1; -;
 DR EMBL; AY029487; AAK33100.1; JOINED.
 DR EMBL; AY029488; AAK33100.1; JOINED.
 DR EMBL; AY029489; AAK33100.1; JOINED.
 DR EMBL; AY029490; AAK33100.1; JOINED.
 DR EMBL; AY029491; AAK33100.1; JOINED.
 DR EMBL; AY029492; AAK33100.1; JOINED.
 DR EMBL; AY029493; AAK33100.1; JOINED.
 DR EMBL; AY029494; AAK33100.1; JOINED.
 DR EMBL; AY029495; AAK33100.1; JOINED.
 DR EMBL; AY029496; AAK33100.1; JOINED.
 DR EMBL; AY029497; AAK33100.1; JOINED.
 DR EMBL; AY029498; AAK33100.1; JOINED.
 DR EMBL; AY029499; AAK33100.1; JOINED.
 DR EMBL; AY029500; AAK33100.1; JOINED.
 DR EMBL; AY029501; AAK33100.1; JOINED.
 DR EMBL; AY029502; AAK33100.1; JOINED.
 DR EMBL; AY029503; AAK33100.1; JOINED.
 DR EMBL; AB011138; BAA25492.1; -;
 DR Genbank; HGNC:13547; ATP10C.
 DR MIM; 605855; -;
 DR MIM; 105830; -;
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR InterPro; IPR001454; Hlgase/hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.


```

FT CONFLICT 2127 2129 GIE -> VSA (IN REF. 4).
FT CONFLICT 2281 2281 G -> E (IN REF. 5).
FT CONFLICT 2413 2413 T -> R (IN REF. 5).
FT CONFLICT 2567 2567 K -> E (IN REF. 5).
SQ SEQUENCE 3174 AA; 360299 MW; FD2AB895B4A7149E CRC64;

Query Match
Best Local Similarity 64.0%; Score 32; DB 1; Length 3174;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| :| | | |
Db 242 ENIVPEGYDF 251

RESULT 15
HIS9_THEME STANDARD; PRT; 233 AA.
AC Q9WZL1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable histidinol-phosphatase (EC 3.1.3.15) (HolPase).
GN HISK OR TM0804.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 393:323-329(1999).
CC -1- CATALYTIC ACTIVITY: L-histidinol-phosphate + H(2)O = L-histidinol
CC + phosphate.
CC -1- PATHWAY: Histidine biosynthesis; eighth step.
CC -1- SIMILARITY: BELONGS TO THE PHP FAMILY OF HYDROLASE. HISK FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001748; AAD35886.1; -
CC TIGR; TM0804; -
CC InterPro; IPR004013; PHP_C.
CC InterPro; IPR003141; PHP_N.
CC Pfam; PF02231; PHP_N; 1.
CC Pfam; PF02811; PHP_C; 1.
CC SMART; SM00481; POLI1A; 1.
KW Histidine biosynthesis; Hydrolase; Complete proteome.
SQ SEQUENCE 233 AA; 271144 MW; E0CE3D245E0EDA30 CRC64;

Query Match
Best Local Similarity 62.0%; Score 31; DB 1; Length 233;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| :| | | |
Db 77 EAVFPDGYDF 86

RESULT 16
YHAL_CRYPA STANDARD; PRT; 319 AA.
AC P10941;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein 1 in hypovirulence-associated DS-RNA genetic
DE element (Contains: P29 proteinase).
OS Cryphonectria parasitica (Chestnut blight fungus) (Endothia
OS parasitica).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Diaporthales; Valsaceae; Cryphonectria.
OX NCBI_TaxID=5116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EP713;
RX MEDLINE=89251594; PubMed=2721496;
RA Rae B.P., Hillman B.I., Tartaglia J., Nuss D.L.;
RT "Characterization of double-stranded RNA genetic elements associated
RT with biological control of chestnut blight: organization of terminal
RT domains and identification of gene products."
RL EMBO J. 8:657-663(1989).
CC -1- MISCELLANEOUS: DOUBLE-STRANDED RNA GENETIC ELEMENTS ARE ASSOCIATED
CC WITH BIOLOGICAL CONTROL OF THE FUNGAL DISEASE CHESTNUT BLIGHT.
CC THIS DS-RNA ARE ASSOCIATED WITH HYPOVIRULENCE. THEY ARE LOCALIZED
CC IN THE CYTOPLASM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C7.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14524; CAA32666.1; -
CC MEROPS; C07.001; -
CC InterPro; IPR002704; Peptidase_C7.
CC Pfam; PF01830; Peptidase_C7; 1.
CC ProDom; PD040949; Peptidase_C7; 1.
KW Hypothetical protein; Hydrolase; Thiol protease.
SQ SEQUENCE 319 AA; 35443 MW; A1F5F77F9AEC7A CRC64;

Query Match
Best Local Similarity 62.0%; Score 31; DB 1; Length 319;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
Db 31 EEVVPAG 37

RESULT 17
ACDM_HUMAN STANDARD; PRT; 421 AA.
AC P11310; Q9NYF1;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
DE (EC 1.3.99.3) (MCAD).
GN ACADM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231952; PubMed=3035565;
RA Kelly D.P., Kim J.-J.P., Billadello J.J., Hainline B.E., Chu T.W.,

```

- RT Strauss A.W.: "Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and its expression in enzyme-deficient human tissue.";
 Proc. Natl. Acad. Sci. U.S.A. 84:4068-4072(1987).
 [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE-Colon;
 RA Sun F., Wang Y., Block G.D.;
 "Medium-chain acyl-CoA dehydrogenase.";
 Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RA Strausberg R.;
 Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
 [4]
 RN SEQUENCE OF 314-342 FROM N.A., AND VARIANT MCAD GLU-329.
 RX MEDLINE-90365752; PubMed-2393404;
 RA Matsubara Y., Narisawa K., Miyabayashi S., Tada K., Coates P.M.,
 Bachmann C., Elsas L.J. II, Pollitt R.J., Rhead W.J., Roe C.R.;
 "Identification of a common mutation in patients with medium-chain
 acyl-CoA dehydrogenase deficiency.";
 Biochem. Biophys. Res. Commun. 171:498-505(1990).
 [5]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE-96420477; PubMed-8823176;
 RA Lee H.J., Wang M., Paschke R., Nandy A., Ghisla S., Kim J.J.;
 "Crystal structures of the wild type and the Glu376Gly/Thr255Glu
 mutant of human medium-chain acyl-CoA dehydrogenase: Influence of the
 location of the catalytic base on substrate specificity.";
 Biochemistry 35:12412-12420(1996).
 [6]
 RN REVIEW ON VARIANTS MCAD.
 RX MEDLINE-93250819; PubMed-1363805;
 RA Tanaka K., Yokota I., Coates P.M., Strauss A.W., Kelly D.P.,
 Zhang Z.F., Gregersen N., Andresen B.S., Matsubara Y., Curtis D.,
 Chen Y.-T.;
 "Mutations in the medium chain acyl-CoA dehydrogenase (MCAD) gene.";
 Hum. Mutat. 1:271-279(1992).
 [7]
 RN VARIANT MCAD GLU-329.
 RX MEDLINE-90368980; PubMed-2394825;
 RA Yokota I., Indo Y., Coates P.M., Tanaka K.;
 "Molecular basis of medium chain acyl-coenzyme A dehydrogenase
 deficiency. An A to G transition at position 985 that causes a
 lysine-304 to glutamate substitution in the mature protein is the
 single prevalent mutation.";
 J. Clin. Invest. 86:1000-1003(1990).
 [8]
 RN VARIANT MCAD GLU-329.
 RX MEDLINE-91067682; PubMed-2251268;
 RA Kelly D.P., Whelan A.J., Ogden M.L., Alpers R., Zhang Z.F., Bellus G.,
 Gregersen N., Dorland L., Strauss A.W.;
 "Molecular characterization of inherited medium-chain acyl-CoA
 dehydrogenase deficiency.";
 Proc. Natl. Acad. Sci. U.S.A. 87:9236-9240(1990).
 [9]
 RN VARIANTS MCAD ILE-149; ARG-244; ARG-267 AND THR-375.
 RX MEDLINE-92081773; PubMed-1684086;
 RA Yokota I., Coates P.M., Hale D.E., Rinaldo P., Tanaka K.;
 "Molecular survey of a prevalent mutation, 985A-to-G transition, and
 identification of five infrequent mutations in the medium-chain acyl-
 CoA dehydrogenase (MCAD) gene in 55 patients with MCAD deficiency.";
 Am. J. Hum. Genet. 49:1280-1291(1991).
 [10]
 RN VARIANT MCAD GLU-329.
 RX MEDLINE-91224627; PubMed-1902818;
 RA Gregersen N., Andresen B.S., Bross P., Winter V., Ruediger N.,
 Engst S., Christensen E., Kelly D., Strauss A.W., Koelvræ S.,
 Bolund L., Ghisla S.;
 "Molecular characterization of medium-chain acyl-CoA dehydrogenase
 (MCAD) deficiency: identification of a lys329 to glu mutation in the
 MCAD gene, and expression of inactive mutant enzyme protein in E.
 coli.";
 Hum. Genet. 86:545-551(1991).
 [11]
 RN VARIANT MCAD GLU-329 FREQUENCY.
 RX MEDLINE-91109464; PubMed-1671131;
 RA Blakemore A.I., Singleton H., Pollitt R.J., Engel P.C., Koelvræ S.,
 Gregersen N., Curtis D.;
 "Frequency of the G985 MCAD mutation in the general population.";
 Lancet 337:298-299(1991).
 [12]
 RN VARIANTS MCAD THR-326 AND ARG-336.
 RX MEDLINE-94256509; PubMed-8198141;
 RA Andresen B.S., Jensen T.G., Bross P., Knudsen I., Winter V.,
 Koelvræ S., Bolund L., Ding J.-H., Chen Y.-T., van Hove J.L.K.,
 Curtis D., Yokota I., Tanaka K., Kim J.-J.P., Gregersen N.;
 "Disease-causing mutations in exon 11 of the medium-chain acyl-CoA
 dehydrogenase gene.";
 Am. J. Hum. Genet. 54:975-988(1994).
 [13]
 RN VARIANT MCAD 115-GLY-CYS-116 DEL.
 RX MEDLINE-95327425; PubMed-7603790;
 RA Ziadeh R., Hofman E.P., Finegold D.N., Hoop R.C., Brackett J.C.,
 Strauss A.W., Naylor E.W.;
 "Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania:
 neonatal screening shows high incidence and unexpected mutation
 frequencies.";
 Pediatr. Res. 37:675-678(1995).
 [14]
 RN VARIANT MCAD ARG-195.
 RX MEDLINE-95015014; PubMed-7929823;
 RA Brackett J.C., Sims H.F., Steiner R.D., Nunge M., Zimmerman E.M.,
 Demartini B., Rinaldo P., Slaughter R., Strauss A.W.;
 "A novel mutation in medium chain acyl-CoA dehydrogenase causes sudden
 neonatal death.";
 J. Clin. Invest. 94:1477-1483(1994).
 [15]
 RN VARIANT MCAD ALA-193.
 RX MEDLINE-97301766; PubMed-9158144;
 RA Andresen B.S., Bross P., Udvari S., Kirk J., Gray G., Knöch S.,
 Chamois N., Knudsen I., Winter V., Wilcken B., Yokota I., Hart K.,
 Packman S., Harper J.P., Saudubray J.M., Hale D.E., Bolund L.,
 Koelvræ S., Gregersen N.;
 "The molecular basis of medium-chain acyl-CoA dehydrogenase (MCAD)
 deficiency in compound heterozygous patients: is there correlation
 between genotype and phenotype?";
 Hum. Mol. Genet. 6:695-707(1997).
 [16]
 RN CHARACTERIZATION OF VARIANT MCAD ALA-193.
 RX MEDLINE-99102080; PubMed-9882619;
 RA Kuchler B., Abdel-Ghany A.G., Bross P., Nandy A., Rasched I.,
 Ghisla S.;
 "Biochemical characterization of a variant human medium-chain acyl-CoA
 dehydrogenase with a disease-associated mutation localized in the
 active site.";
 Biochem. J. 337:225-230(1999).
 [17]
 RN VARIANT HIS-67.
 RX PubMed-11349232;
 RA Andresen B.S., Dobrowolski S.F., O'Reilly L., Muenzer J.,
 McCandless S.E., Frazier D.W., Udvari S., Bross P., Knudsen I.,
 Banas R., Chace D.H., Engel P.C., Naylor E.W., Gregersen N.;
 "Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by
 MS/MS-based prospective screening of newborns differ from those
 observed in patients with clinical symptoms: identification and
 characterization of a new, prevalent mutation that results in mild
 MCAD deficiency.";
 Am. J. Hum. Genet. 68:1408-1418(2001).
 [18]
 RN VARIANT MCAD LEU-245.
 RX PubMed-11409868;
 RA Zschöck J., Schulze A., Lindner M., Fiesel S., Olgemöller K.,
 Hoffmann G.F., Penzien J., Rüter J.P.N., Wanders R.J.A.,
 Mayatepek E.;

RT "Molecular and functional characterization of mild MCAD deficiency.";
RL Hum. Genet. 108:404-408(2001).
RN [19]
RP VARIANT THR-281.
RX PubMed=11486912;
RA Albers S., Levy H.L., Irons M., Strauss A.W., Marsden D.;
RT "Compound heterozygosity in four asymptomatic siblings with
medium-chain acyl-CoA dehydrogenase deficiency.";
RL J. Inher. Metab. Dis. 24:417-418(2001).
CC -1- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO 16.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced ETF.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first step.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- DISEASE: DEFECTS IN ACADM ARE THE CAUSE OF MCAD DEFICIENCY, AN AUTOSOMAL RECESSIVE DISEASE WHICH CAUSES FASTING HYPOLYCEMIA, HEPATIC DYSFUNCTION, AND ENCEPHALOPATHY, OFTEN RESULTING IN DEATH IN INFANCY. THE DISEASE FREQUENCY IS ONE IN 13000.
CC -1- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-CoA DEHYDROGENASES OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN TISSUES.
CC -1- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M91432; AAA59567.1; JOINED.
DR EMBL; M91421; AAA59567.1; JOINED.
DR EMBL; M91422; AAA59567.1; JOINED.

Query Match 62.0%; Score 31; DB 1; Length 421;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 EEVVPXGXDY 10
DB 58 EEIIPVAAEY 67
||:|:|:|

RESULT 18
ACDM_PIG STANDARD; PRT; 421 AA.
ID ACDM_PIG
AC P41367;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor (EC 1.3.99.3) (MCAD).
GN ACADM.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki H., Kimura M., Ito T., Murakami Y., Hamasima N., Yasue H.;
RL Submitted (Nov-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93361479; PubMed=8356049;
RA Kim J.-J.P., Wang M., Paschke R.;
RT "Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:7523-7527(1993).
CC -1- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO 16.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced ETF.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first step.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-CoA DEHYDROGENASES OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN TISSUES.
CC -1- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40845; AAA83759.1; -.
DR PDB; 3MDD; 30-SEP-94.
DR PDB; 3MDE; 30-SEP-94.
DR InterPro: IPR001552; Acyl-CoA_dh.
DR Pfam: PF00441; Acyl-CoA_dh; 1.
DR Pfam: PF02770; Acyl-CoA_dh_M; 1.
DR Pfam: PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE: PS00072; ACYL_COA_DH_1; 1.
DR PROSITE: PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW Mitochondrion; Transit peptide; 3D-structure.
FT TRANSIT 1 25 MITOCHONDRION.
FT CHAIN 26 421 ACYL-CoA DEHYDROGENASE, MEDIUM-CHAIN SPECIFIC.
FT ACT_SITE 193 193 FORMS A HYDROGEN-BOND WITH THE FLAVIN N(5) OF THE FAD COFACTOR (BY SIMILARITY).
FT ACT_SITE 401 401 BASE.
FT CONFLICT 40 40 K -> E (IN REF. 2).
FT CONFLICT 283 283 S -> P (IN REF. 2).
FT CONFLICT 305 305 G -> E (IN REF. 2).
FT CONFLICT 331 331 D -> E (IN REF. 2).
SQ SEQUENCE 421 AA; 46358 MW; 27347B0C99F2B7C3 CRC64;
Query Match 62.0%; Score 31; DB 1; Length 421;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 EEVVPXGXDY 10
DB 58 EEIIPVAAEY 67
||:|:|:|

RESULT 19
TR2M_ERWHE STANDARD; PRT; 562 AA.
ID TR2M_ERWHE
AC Q47861;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN IAM.
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=gv. gyropsophila / PD713;
RA Ophir Y., Kunik T., Manulis S., Lichter A., Barash I., Gafni Y.;
RL Submitted (Nov-1995) to the EMBL/GenBank/DBJ databases.

CC -|- CATALYTIC ACTIVITY: L-tryptophan + O(2) = indole-3-acetamide +
 CC CO(2) + H(2)O.
 CC -|- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF AUXINS FROM TRYPTOPHAN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33867; AAC17187.1; -
 CC InterPro: IPR002937; Amino_Oxidase.
 CC InterPro: IPR00205; NAD_binding.
 CC Pfam: PF01593; Amino_Oxidase; 1.
 CC Oxidoreductase; Monooxygenase; Auxin biosynthesis.
 CC SEQUENCE 562 AA; 62513 MW; 621E7AA389DF886 CRC64;
 CC
 CC Query Match 62.0%; Score 31; DB 1; Length 562;
 CC Best Local Similarity 62.5%; Pred. No. 63;
 CC Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 3 VVPXGXDY 10
 CC :|||
 CC DB 456 LVPAGADY 463
 CC
 CC RESULT 20
 CC PLAS_DAUCA STANDARD; PRT; 97 AA.
 CC ID P20422;
 CC DT 01-FEB-1991 (Rel. 17, Created)
 CC DT 01-FEB-1991 (Rel. 17, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Plastocyanin.
 CC GN PETE.
 CC OS Daucus carota (Carrot).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 CC OX NCBI_TaxID=4039;
 CC RN [1]
 CC RP SEQUENCE.
 CC RA Shoji A., Yoshizaki F., Karahashi A., Sugimura Y., Shimokoriyama M.;
 CC RT "Complete amino acid sequence of plastocyanin from Daucus carota.";
 CC RL Selkagaku 57:1036-1036(1985).
 CC CC -|- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC CC -|- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC CC -|- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC PIR: JWO011; JWO011.
 CC DR HSSP: P17341; IPLB.
 CC DR InterPro: IPR000923; BlueCu1.
 CC DR InterPro: IPR001235; Copper blue.
 CC DR Pfam: PF00127; copper-binding; 1.
 CC DR PRINTS: PR00156; COPPERBLUE.
 CC DR PRODOM: PD001235; Copper_blue; 1.
 CC DR PROSITE: PS00196; COPPER_BLUE; 1.
 CC KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane.
 CC FT DOMAIN 1 97 PLASTOCYANIN-LIKE.
 CC FT METAL 37 37 COPPER (BY SIMILARITY).
 CC FT METAL 82 82 COPPER (BY SIMILARITY).
 CC FT METAL 85 85 COPPER (BY SIMILARITY).
 CC FT METAL 90 90 COPPER (BY SIMILARITY).
 CC SEQUENCE 97 AA; 10181 MW; B15DE6B8428F72D4 CRC64;
 CC
 CC Query Match 60.0%; Score 30; DB 1; Length 97;
 CC Best Local Similarity 54.5%; Pred. No. 16;
 CC Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 EEVVPXGXDS 11

Db 43 EDEVPGVDVS 53

:|:|:|:|

RESULT 21

CC TRNK_HUMAN STANDARD; PRT; 121 AA.
 CC ID TRNK_HUMAN
 CC AC Q9UHF0;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Neurokinin B precursor (NKB) (Neuromedin K) (ZNEUROK1).
 CC GN TAC3.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,
 CC RL O'Hara P.;
 CC RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Placenta;
 CC RX MEDLINE=20322570; PubMed=10866201;
 CC RA Page N.M., Woods R.J., Gardiner S.M., Lomthiasong K., Gladwell R.T.,
 CC RA Butlin D.J., Manyonda I.T., Lowry P.J.;
 CC RT "Excessive placental neurokinin B secretion during the third trimester
 CC causes pre-eclampsia.";
 CC RL Nature 405:797-800(2000).
 CC CC -|- FUNCTION: TACHIKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES (By similarity).
 CC CC -|- SUBCELLULAR LOCATION: Secreted.
 CC CC -|- DEVELOPMENTAL STAGE: In pregnancy, the expression of NKB is
 CC confined to the outer syncytiotrophoblast of the placenta,
 CC significant concentrations of NKB can be detected in plasma as
 CC early as week 9, and plasma concentrations of NKB are grossly
 CC elevated in pregnancy-induced hypertension and pre-eclampsia.
 CC CC -|- SIMILARITY: BELONGS TO THE TACHIKININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF186112; AAF01430.1; -
 CC EMBL: AF216586; AAF76980.1; -
 CC Genew: HGNC:11521; TAC3.
 CC MIM: 162330;
 CC InterPro: IPR003635; Neurokinin.
 CC InterPro: IPR002040; Tachykinin.
 CC PRODOM: PD020370; Neurokinin; 1.
 CC DR PROSITE: PS00267; TACHYKININ; 1.
 CC DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
 CC KW Amidation; Signal.
 CC FT SIGNAL 1 16 POTENTIAL.
 CC FT PROPEP 17 78 BY SIMILARITY.
 CC FT PEPTIDE 81 90 NEUROKININ B.
 CC FT PROPEP 94 121 BY SIMILARITY.
 CC FT MOD_RES 90 90 AMIDATION (G-91 PROVIDE AMIDE GROUP) (BY
 CC SIMILARITY).
 CC SEQUENCE 121 AA; 13438 MW; 14C9AFE2EE9DECA CRC64;
 CC
 CC Query Match 60.0%; Score 30; DB 1; Length 121;
 CC Best Local Similarity 85.7%; Pred. No. 20;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 28 EEVPPGG 34

RESULT 22

ID HES3_RAT STANDARD; PRT; 175 AA.
 AC Q04667; 1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Transcription factor HES-3 (Hairy and enhancer of split 3).
 GN HES3 OR HES-3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=Sprague-Dawley; TISSUE=Brain;
 MEDLINE=94040724; PubMed=1340473;
 RA Sasai Y., Kageyama R., Tagawa Y., Shigemoto R., Nakanishi S.;
 RT "Two mammalian helix-loop-helix factors structurally related to
 Drosophila hairy and enhancer of split.";
 RL Genes Dev. 6:2620-2634(1992).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
 CC PROTEIN FOR THEIR TRANSCRIPTION.
 CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
 CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN PURKINJE CELLS.
 CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 CC HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
 CC RATHER THAN THE CANONICAL E-BOX (CANNIG).
 CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
 CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 CC HAIRY-RELATED PROTEINS
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D13418; BAA02683.1; -
 DR PIR: S36749; S36749.
 DR TRANSFAC: T01651; -
 DR InterPro: IPR001092; HLB_baslc.
 DR Pfam: PF00010; HLB; 1.
 DR SMART: SM00353; HLB; 1.
 DR PROSITE: PS00038; HLB_1; 1.
 DR PROSITE: PS00888; HLB_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Repressor.
 FT DNA_BIND 1 6
 FT DOMAIN 7 50
 FT DOMAIN 108 167
 FT DOMAIN 172 175
 FT WRPW MOTIF (REQUIRED FOR ACTIVITY)
 FT (BY SIMILARITY)
 SQ SEQUENCE 175 AA; 19137 MW; D490663869155CB3 CRC64;
 Query Match 60.0%; Score 30; DB 1; Length 175;
 Best Local Similarity 62.5%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VVPXGXDY 10
 Db 58 LVPSGVGY 65

RESULT 23

CTC_BACSU STANDARD; PRT; 203 AA.
 AC F14194;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE General stress protein etc.
 GN Ctc.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=168;
 RX MEDLINE=96051385; PubMed=7584024;
 RA Ogasawara N., Nakai S., Yoshikawa H.;
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus
 RT subtilis chromosome containing the replication origin.";
 RL DNA Res. 1:1-14(1994).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Broussier R., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Puigc P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 RN [3]
 RC SEQUENCE OF 1-185 FROM N.A.
 RP MEDLINE=90066361; PubMed=2555671;
 RA Nilsson D., Hove-Jensen B., Arnvig K.;
 RT "Primary structure of the tms and prs genes of Bacillus subtilis.";
 RL Mol. Gen. Genet. 218:565-571(1989).
 RN [4]
 RC SEQUENCE OF 1-13.
 RP STRAIN=168 / IS58;
 RX MEDLINE=94282319; PubMed=8012595;
 RA Voelker U., Engelmann S., Maul B., Riethdorf S., Voelker A.,
 RA Schmid R., Mach H., Hecker M.;
 RT "Analysis of the induction of general stress proteins of Bacillus
 RT subtilis.";
 RL Microbiology 140:741-752(1994).
 CC -1- FUNCTION: NOT KNOWN.

CC -|- DEVELOPMENTAL STAGE: EXPRESSED AT THE END OF EXPONENTIAL GROWTH
 CC UNDER CONDITIONS IN WHICH THE ENZYMES OF THE TCA CYCLE ARE
 CC REPRESSED.
 CC -|- INDUCTION: BY HEAT SHOCK, SALT STRESS, OXIDATIVE STRESS, GLUCOSE
 CC LIMITATION AND OXYGEN LIMITATION.
 CC -|- SIMILARITY: BELONGS TO THE L25P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D26185; BAA05287.1; -;
 DR EMBL; Z99104; CAB11828.1; -;
 DR EMBL; X16518; CAA34524.1; -;
 DR PIR; S05373; S05373;
 DR Subtilist; BG10115; etc.
 DR InterPro; IPR001021; Ribosomal_L25;
 DR Pfam; PF01386; Ribosomal_L25p; 1;
 DR ProDom; PD012503; Ribosomal_L25; 1;
 DR TIGRFAMs; TIGR00731; ctc.TL5; 1;
 DR Heat shock; Complete proteome.
 FT INIT_MET 0
 FT VARIANT 1 1 A -> R (IN STRAIN IS58).
 FT VARIANT 9 9 T -> Q (IN STRAIN IS58).
 FT VARIANT 12 12 T -> I (IN STRAIN IS58).
 SQ SEQUENCE 203 AA; 21924 MW; 5130A04D0E86247 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 203;
 Best Local Similarity 62.5%; Pred. No. 35;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGDYS 11
 : | | | |
 DB 157 LPAGGDYS 164

RESULT 24

ARAD_ECOLI STANDARD; PRT; 231 AA.
 ID ARAD_ECOLI
 AC P08203;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) (Phosphoribulose
 DE isomerase).
 GN ARAD OR B0061.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RX SEQUENCE FROM N.A.
 RC STRAIN-B;
 RX MEDLINE-87163495; PubMed-3549454;
 RA Lee N., Gielow W., Martin R., Hamilton E., Fowler A.;
 RT "The organization of the arabad operon of Escherichia coli.";
 RL Gene 47:231-244 (1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-91083835; PubMed-2261080;
 RA Chen H., Sun Y., Stark T., Beattie W., Moses R.E.;
 RT "Nucleotide sequence and deletion analysis of the polB gene of
 RT Escherichia coli.";
 RL DNA Cell Biol. 9:631-635 (1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-91067495; PubMed-2251150;

RA Mineno J., Fukui H., Ishino Y., Kato I., Shinagawa H.;
 RT "Nucleotide sequence of the arad gene of Escherichia coli K12
 RT encoding the L-ribulose 5-phosphate 4-epimerase.";
 RL Nucleic Acids Res. 18:6722-6722(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-92334977; PubMed-1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474 (1997).
 RN [6]
 RP SEQUENCE OF 222-231 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-91017565; PubMed-2217198;
 RA Bonner C.A., Hays S., McEntee K., Goodman M.F.;
 RT "DNA polymerase II is encoded by the DNA damage-inducible dna gene
 RT of Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:7663-7667(1990).
 RN [7]
 RP SEQUENCE OF 158-231 FROM N.A.
 RC STRAIN-K12 / W3110;
 RX MEDLINE-91238699; PubMed-2034216;
 RA Iwasaki H., Ishino Y., Itoh H., Nakata A., Shinagawa H.;
 RT "Escherichia coli DNA polymerase II is homologous to alpha-like DNA
 RT polymerases.";
 RL Mol. Gen. Genet. 226:24-33(1991).
 CC -|- CATALYTIC ACTIVITY: L-ribulose 5-phosphate = D-xylulose 5-
 CC phosphate.
 CC -|- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
 CC -|- PATHWAY: L-arabinose catabolism; third step.
 CC -|- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ARAD/FUCA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; M15263; AAA23464.1; -;
 EMBL; M35371; -; NOT_ANNOTATED_CDS.
 EMBL; M62646; AAA24405.1; -;
 EMBL; D10483; BAA01332.1; -;
 EMBL; AE000116; AAC73172.1; -;
 EMBL; M37727; AAA23683.1; -;
 EMBL; M38283; AAA63763.1; -;
 EMBL; X56048; CAA39519.1; -;
 PIR; D29022; ISECP4.
 PIR; S13593; ISECK4.
 PIR; A36236; A36236.
 PIR; S40577; S40577.
 ECO2DBASE; G028.1; 6TH EDITION.
 Ecogene; EG10055; arad.
 InterPro; IPR001303; Aldolase_II_N.
 InterPro; IPR004661; Arad.
 Pfam; PF00596; Aldolase_II; 1.
 TIGRFAMs; TIGR00760; arad; 1.
 KW Arabinose catabolism; isomerase; zinc; Complete proteome.

```
FT METAL 76 76 ZINC (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
FT METAL 97 97 ZINC (BY SIMILARITY).
FT METAL 171 171 ZINC (BY SIMILARITY).
FT VARIANT 50 50 V -> I.
FT VARIANT 70 70 T -> A.
FT VARIANT 216 216 D -> N.
SQ SEQUENCE 231 AA; 25519 MW; 1753F75958332163 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 231;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDXYS 11
Db 41 IKPSGVDYS 49

RESULT 25
ARAD_SALTY
ID ARAD_SALTY STANDARD; PRT; 231 AA.
AC P06190;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) (Phosphoribulose
isomerase).
DE ARAD OR STM0101.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=85232046; PubMed=3891514;
RA Lin H.-C., Lei S.-P., Studnicka G., Wilcox G.;
RT "The arabid operon of Salmonella typhimurium LT2. III. Nucleotide
sequence of arab and its flanking regions, and primary structure of
its product, L-ribulose-5-phosphate 4-epimerase."
RL Gene 34:129-134(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SCSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856(2001).
CC -!- CATALYTIC ACTIVITY: L-ribulose 5-phosphate = D-xylulose 5-
phosphate.
CC -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC -!- PATHWAY: L-arabinose catabolism: third step.
CC -!- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ARAD/FUCA
SUBFAMILY.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
frameshift in position 202.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M11047; AAA27025.1; ALT_FRAME.
CC DR EMBL; AE008698; AAL19065.1; -.
CC DR PIR; A24986; ISEB4T.
```

```
DR StyGene; SG10015; arad.
DR InterPro; IPR001303; Aldolase_II_N.
DR InterPro; IPR004661; Arad.
DR Pfam; PF00596; Aldolase_II; 1.
DR TIGRFAMS; TIGR00760; arad; 1.
KW Arabinose catabolism; isomerase; zinc; Complete proteome.
FT METAL 76 76 ZINC (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
FT METAL 97 97 ZINC (BY SIMILARITY).
FT METAL 171 171 ZINC (BY SIMILARITY).
SQ SEQUENCE 231 AA; 25531 MW; DA473505739284F9 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 231;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDXYS 11
Db 41 IKPSGVDYS 49
```

Search completed: June 10, 2003, 13:40:24
Job time : 4.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-45
Perfect score: 50
Sequence: 1 EFVVPXGXDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

- Database : SPTREMBL.21:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	36	72.0	298	2	052367	052367	rhizobium t
2	36	72.0	363	17	030260	030260	archaeoglob
3	35	70.0	78	6	09XST4	09XST4	canis fami
4	35	70.0	143	5	09VSY8	09VSY8	drosophila
5	35	70.0	217	4	000404	000404	homo sapien
6	35	70.0	290	16	08U7J0	08U7J0	agrobacteri
7	35	70.0	299	4	09UEE9	09UEE9	homo sapien
8	35	70.0	587	16	09JZP8	09JZP8	neisseria m
9	35	70.0	692	5	09V190	09V190	drosophila
10	35	70.0	906	10	09ZSY4	09ZSY4	arabidopsis
11	35	70.0	908	10	09FJK8	09FJK8	arabidopsis
12	35	70.0	908	10	08W4J9	08W4J9	arabidopsis
13	35	70.0	908	10	09ZSY3	09ZSY3	arabidopsis
14	35	70.0	909	10	09W5A1	09W5A1	arabidopsis
15	35	70.0	1063	16	08RG86	08RG86	fusobacteri
16	34	68.0	156	3	012479	012479	saccharomyc

08VY88 arabidopsis
092u66 rhizobium m
09c610 arabidopsis
09a382 caulobacter
09v9t5 sulfolobus
09v9t6 drosophila
08t62 methanopyru
092bc5 listeria in
08v6u8 listeria mo
08zaq9 yersinia pe
065890 cycloclolla
040129 lycopersico
08xp8 clostridium
09d876 mus musculu
09yfi3 aeropyrum p
08r126 mus musculu
08vd18 mus musculu
09x215 raiistonla s
09ddj4 halichoeres
099x15 streptococc
09elx6 cercopithec
08y111 brucella me
09kt32 vibrio chol
017704 caenorhabdi
098k29 rhizobium l
09crg3 mus musculu
08tng3 methanosarc
09abb8 oryza sativ
09gq04 eriocheir s
098362 human immun
09jma5 mus musculu
09vix6 drosophila
09vnm9 drosophila
070565 mus musculu
088271 mus musculu
099td4 staphylococ
09np60 halobacteri
08yyj8 anabaena sp
09d8m6 mus musculu
09cx34 mus musculu
09cre7 mus musculu
029920 archaeoglob
029451 archaeoglob
08ty23 methanopyru
09xv00 rhizobium m
09xvk4 caenorhabdi
052680 escherichia
08tf64 trichoderma
08u2a4 pyrococcus
09cin1 lactococcus
09hnp8 halobacteri
045063 caenorhabdi
095p46 carcinus ma
091kl4 zea mays (m
094f60 dichanthell
09u6a3 callinectes
094dz5 oryza sativ
052673 escherichia
052666 escherichia

ALIGNMENTS

RESULT 1			
052367	PRELIMINARY:	PRT;	298 AA.
ID	052367		
AC	052367:		
DT	01-JUN-1998	(TREMBLrel. 06, Created)	
DT	01-JUN-1998	(TREMBLrel. 06, Last sequence update)	
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)	
DE	Aryl-alcohol dehydrogenase homolog (Fragment).		
GN	XYLB1.		

OS Rhizobium tropici.
 OG Plasmid pRtrCFN299a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=398;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CFN299;
 RA Rosenblueth M., Hynes M.F., Martinez-Romero E.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- COFACTOR: ZINC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 DR EMBL; AF036920; AAC04779.1; -.
 DR HSP; P07846; 1SDG.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; Adh_zn_family.
 DR Pfam; PF00107; adh_zinc; 1.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Plasmid; Zinc.
 FT NON_TER 298
 SQ SEQUENCE 298 AA; 31092 MW; 49B2F8117C3AE87 CRC64;
 Query Match 72.0%; Score 36; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPGXDYS 11
 I:| | | |
 Db 250 EIIPEGADFS 259
 RESULT 2
 O30260
 ID O30260 PRELIMINARY; PRT; 363 AA.
 AC O30260;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein AF2411.
 GN AF2411.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Richardson D.L., Kerlavage A.R., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91255.1; -.
 DR TIGR; AF2411; -.
 DR InterPro; IPR002103; Bac_luciferase.
 DR Pfam; PF00296; bac_luciferase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 363 AA; 41736 MW; 0B976AE788F4803 CRC64;
 Query Match 72.08; Score 36; DB 17; Length 363;
 Best Local Similarity 54.5%; Pred. No. 15;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGDYS 11
 I:| | | |
 Db 120 ENIVPYGIDFS 130
 RESULT 3
 Q9XST4
 ID Q9XST4 PRELIMINARY; PRT; 78 AA.
 AC Q9XST4;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE P97 homologous protein (Fragment).
 GN P97.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYROID.
 RX MEDLINE=20422104; PubMed=10964405;
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
 RA Christophe D.;
 RT "A method for the large-scale cloning of nuclear proteins and nuclear
 RT targeting sequences on a functional basis.";
 RL Anal. Biochem. 284:231-239(2000).
 DR EMBL; AJ388531; CAB46833.1; -.
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8895 MW; B62486313555FBA1 CRC64;
 Query Match 70.0%; Score 35; DB 6; Length 78;
 Best Local Similarity 54.5%; Pred. No. 4.4;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGDYS 11
 I:| | | |
 Db 16 EDVPSGGEYS 26
 RESULT 4
 Q9VSY8
 ID Q9VSY8 PRELIMINARY; PRT; 143 AA.
 AC Q9VSY8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CG3911 protein.
 GN CG3911.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003552; AAF50270.1; -;
 DR FLYBase: FBgn0035992; CG3911.
 SQ SEQUENCE 143 AA; 16471 MW; 1DC346DC22C02AA2 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 143;
 Best Local Similarity 66.7%; Pred. No. 8.7;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXD 9
 |||||
 Db 135 EEVVPAGED 143

RESULT 5

ID 000404 PRELIMINARY; PRT; 217 AA.

AC 000404;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE P97 homologous protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX TISSUE=PLACENTA;
 RX MEDLINE=97160586; PubMed=9006920;
 RA Nobukuni T., Kobayashi M., Oomori A., Ichinose S., Iwanaga T.,
 RA Takahashi I., Hashimoto K., Hattori S., Kaibuchi K., Miyata Y.;
 RT "An Alu-linked repetitive sequence corresponding to 280 amino acids is
 expressed in a novel bovine protein, but not in its human homologue.";
 RL J. Biol. Chem. 272:2801-2807(1997).
 DR EMBL: D85939; BAA20069.1; -;
 SQ SEQUENCE 217 AA; 24061 MW; E404BA2E35497828 CRC64;

Query Match 70.0%; Score 35; DB 4; Length 217;
 Best Local Similarity 54.5%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 |::|::|
 Db 16 EDVVPSSGEYS 26

RESULT 6

Q08U7J0

ID Q08U7J0 PRELIMINARY; PRT; 290 AA.
 AC Q08U7J0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 6-O-methylguanine-DNA methyltransferase.
 GN ADA OR ATU4459 OR AGR.L818.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McLelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak-C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Ramm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Ouello B., Gindman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmiel K., Gordon J., Vaudin M., Hartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009374; AAL45253.1; -;
 DR EMBL: AE008240; AAK8982.1; -;
 KW Methyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 290 AA; 31587 MW; B626592EF519977F CRC64;

Query Match 70.0%; Score 35; DB 16; Length 290;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 |::|::|
 Db 9 EDITPGSDY 18

RESULT 7

ID Q09UE9 PRELIMINARY; PRT; 299 AA.

AC Q09UE9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BCNT protein (CRANIOFACIAL development protein 1).
 GN BCNT.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.

RX MEDLINE=98267221; PubMed=9602175;
 RA Takahashi I., Nobukuni T., Oomori H., Kobayashi M., Tanaka S.,
 RA Ohshima K., Okada N., Masui T., Hashimoto K., Iwashita S.;
 RT "Existence of a bovine LINE repetitive insert that appears in the cdna
 of bovine protein BCNT in ruminant, but not in human, genomes.";

```

RL Gene 211:387-394(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB009285; BAA31867.1; -
DR EMBL; BC000991; AAH00991.1; -
SQ SEQUENCE 299 AA; 33593 MW; F4A9E928B669451A CRC64;

Query Match 70.0%; Score 35; DB 4; Length 299;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDS 11
Db 16 EYVPSGGEYS 26

RESULT 8
RQ9JZP8 PRELIMINARY; PRT; 587 AA.
ID Q9JZP8;
AC Q9JZP8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Succinate dehydrogenase, flavoprotein subunit.
GN NM00950.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OC NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WC58 / SEROGROUP B;
RC MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey T.R., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58".
RL Science 287:1809-1815(2000).
CC -1-COFACITOR: FAD (BY SIMILARITY).
DR EMBL; AE002446; AAF41356.1; -
DR TIGR; NMB0950; -
DR InterPro; IPR003953; FAD_bind2.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR003952; FRD/SDH_FAD.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR004112; Succ_DH_flav_C.
DR Pfam; PF00890; FAD_binding_2; 1.
DR Pfam; PF02910; succ_DH_flav_C; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRTDASE1.
DR PROSITE; PS00504; FRD_SDH_FAD_BINDING; 1.
DR FAD; Flavoprotein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 587 AA; 64502 MW; 9581701B08069003 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 587;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10
Db 366 EVVVPQGEDY 375

RESULT 9

```

```

Q9VI90
ID Q9VI90 PRELIMINARY; PRT; 692 AA.
AC Q9VI90;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CG10040 protein (Roughened eye).
GN RN CG10040 OR CG14600 OR CG14601.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C.R., Miklos G.L.G.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Flossler K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spiers R., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swierkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC St Pierre S.E., Galindo M.I., Couso J.P., Thor S.;
RT "Control of Drosophila imaginal disc development by rotund and
RT roughened eye: differentially expressed transcripts of the same gene
RT encoding distinct zinc finger proteins."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003672; AAF54035.1; -
DR EMBL; AF395904; AAL59598.1; -
DR HSSP; P08153; 1ZFD.
DR FlyBase; FBgn0037494; rn.
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR Prodom; PD000003; znf_C2H2; 1.
DR SMART; SM00355; znf_C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
DR DNA-binding; Metal-binding; Zinc-finger.

```

SQ SEQUENCE 692 AA; 75774 MW; BD010502BB65042E CRC64;

Query Match 70.0%; Score 35; DB 5; Length 692;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPGXGDYS 11
 I::|||
 Db 26 EIPAGGDYS 35

RESULT 10
 Q9ZSY4 PRELIMINARY; PRT; 906 AA.

AC Q9ZSY4
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein RPP8.
 GN RPP8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LA-ER;
 RA MEDLINE=99030193; PubMed=9811794;
 RX McDowell J.M., Dhandaydham M., Long T.A., Aarts M.G., Goff S.,
 RA Holub E.B., Dangl J.L.;
 RT "Intragenic recombination and diversifying selection contribute to the
 evolution of downy mildew resistance at the RPP8 locus of
 Arabidopsis.";
 RL Plant Cell 10:1861-1874(1998).
 DR EMBL; AF089710; AAC83165.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 1.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 906 AA; 104201 MW; 52905EB143675F8F CRC64;

Query Match 70.0%; Score 35; DB 10; Length 906;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPGXGDY 10
 I::|||
 Db 881 EKLVPGGEDY 890

RESULT 11
 Q9FJK8 PRELIMINARY; PRT; 908 AA.

AC Q9FJK8
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein.
 GN RPP8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99087489; PubMed=9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.

RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 physically assigned pl and TAC clones.";
 RL DNA Res. 5:297-308(1998).
 DR EMBL; AB015468; BAB10695.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104448 MW; 3111991B17239693 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPGXGDY 10
 I::|||
 Db 883 EKLVPGGEDY 892

RESULT 12
 Q8W4J9 PRELIMINARY; PRT; 908 AA.

AC Q8W4J9
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein RPP8.
 GN AT5G43470, MW20.19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H., Yamada K.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY062514; AAL32592.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104649 MW; 4461F553128F3A15 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPGXGDY 10
 I::|||
 Db 883 EKLVPGGEDY 892

RESULT 13
 Q9ZSY3 PRELIMINARY; PRT; 908 AA.

AC Q9ZSY3
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RPP8 (disease resistance protein RPP8).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COL;
 RX MEDLINE=99030193; PubMed=9811794;
 RA McDowell J.M., Dhandaaydham M., Long T.A., Aarts M.G.M., Goff S.,
 RA Holub E.B., Dangl J.L.;
 RT "Intragenic recombination and diversifying selection contribute to the
 RT evolution of downy mildew resistance at the RPP8 locus of
 RT Arabidopsis.";
 RL Plant Cell 10:1861-1874 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63 (2000).
 DR EMBL; AF089711; AAC78631.1; -;
 DR EMBL; AB025638; BAA97426.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104681 MW; 1BFA35BB6B0CB5CD CRC64;

Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 Db 883 EKLVPGGEDY 892
 I::| | | | |
 RESULT 14
 Q9M5A1
 ID Q9M5A1 PRELIMINARY; PRT; 909 AA.
 AC Q9M5A1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Viral resistance protein.
 GN HRT.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DI-17;
 RA Cooley M.B., Pathirana S., Wu H., Kachroo P., Klessig D.F.;
 RT "Members of the Arabidopsis HRT/RPP8 family of resistance genes confer
 RT resistance to both viral and oomycete pathogens.";
 RL Plant Cell 0:0-0 (2000).
 DR EMBL; AF234174; AAF36987.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 909 AA; 105052 MW; 06262B71A2B3037F CRC64;

Query Match 70.0%; Score 35; DB 10; Length 909;
 Best Local Similarity 60.0%; Pred. No. 70;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 Db 884 EKLVPGGEDY 893
 I::| | | | |
 RESULT 15
 Q8RG86
 ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
 AC Q8RG86;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
 GN FN0422.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteriia; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 2586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fonstein M., Kyripides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 2586.";
 RL J. Bacteriol. 184:2005-2018 (2002).
 DR EMBL; AF010554; AAL94625.1; -;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 84;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXDYS 11
 Db 195 EIVPGLNYS 204
 I::| | | | |
 RESULT 16
 Q12479
 ID Q12479 PRELIMINARY; PRT; 156 AA.
 AC Q12479;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ORF YOR013W.
 GN YOR013W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC De haan M., Grivell L.A., Maarse A.C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;
 RA De haan M., Maarse A.C., Grivell L.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;

```

RX MEDLINE=94019318; PubMed=8413243;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RA Sherman F.;
RT "CYC2 encodes a factor involved in mitochondrial import of yeast
RL cytochrome c.";
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
DR EMBL; 274920; CAA99201.1; -
DR EMBL; X87331; CAA60762.1; -
DR SGD; S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 68.0%; Score 34; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVXPXGXDY 10
II:| | |
DB 50 EVVPLGMDY 58

RESULT 17
Q8VY88 PRELIMINARY; PRT; 175 AA.
ID O8VY88
AC O8VY88
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 19.2 kDa protein.
GN ATIG50910.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bann J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072348; AAL61955.1; -
DR InterPro; IPR002110; ANK.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein.
SQ SEQUENCE 175 AA; 19189 MW; 7F72AB1EC82C4190 CRC64;

Query Match 68.0%; Score 34; DB 10; Length 175;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
II::| | |
DB 135 EELLKAGADYS 145

RESULT 18
Q92U66 PRELIMINARY; PRT; 247 AA.
ID Q92U66
AC Q92U66
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

```

```

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein RB1273.
GN RB1273 OR SWB21444.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603646; CAC49673.1; -
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 247 AA; 28930 MW; 2F14F383E66D420 CRC64;

Query Match 68.0%; Score 34; DB 16; Length 247;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
II:| | | |
DB 48 EDVEPRGADY 57

RESULT 19
Q9C6J0 PRELIMINARY; PRT; 257 AA.
ID Q9C6J0
AC Q9C6J0
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 28.6 kDa protein.
GN F8A12.12.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC079284; AAG50930.1; -
DR InterPro; IPR002110; ANK.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein; Repeat.

```

SQ SEQUENCE 257 AA; 28578 MW; 714C7A4387F32B5C CRC64;
 Query Match 68.0%; Score 34; DB 10; Length 257;
 Best Local Similarity 54.5%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDXS 11
 ||:| | || |
 Db 217 EELLKAGADYS 227

 RESULT 20
 Q9A382 PRELIMINARY; PRT; 433 AA.
 AC Q9A382;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Peptidoglycan-binding protein, putative.
 GN CC3322.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 CC Caulobacter.
 OX NCBI_TaxID=155892;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 Uitterback T., Tran A., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005994; AAK25284.1;
 DR HSSP; P41052; 1LTM.
 DR TIGR; CC3322.
 DR InterPro; IPR002477; PG_binding.
 DR Pfam; PF01471; PG_binding_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

 Query Match 68.0%; Score 34; DB 16; Length 433;
 Best Local Similarity 54.5%; Pred. No. 51;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDXS 11
 | : | | | |
 Db 266 EVILPPGFDYS 276

 RESULT 21
 Q96YH5 PRELIMINARY; PRT; 1442 AA.
 AC Q96YH5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein Srt2195.
 GN Srt2195.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 OX NCBI_TaxID=111955;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,

RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AF000989; BAB67302.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1442 AA; 156497 MW; D63EC2C35228121F CRC64;

 Query Match 68.0%; Score 34; DB 17; Length 1442;
 Best Local Similarity 50.0%; Pred. No. 2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDX 10
 ||:| | | |
 Db 863 EEITPGANY 872

 RESULT 22
 Q9V9T6 PRELIMINARY; PRT; 2778 AA.
 AC Q9V9T6; Q9Y0Z7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE FAF protein (BCDNA:LD22582)
 GN FAF OR BCDNA:LD22582 OR CG1945.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers J.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbal J.,
 Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of *Drosophila melanogaster*."
 RN Science 287:2185-2195(2000).
 RP SEQUENCE OF 1090-2778 FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.D., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
 RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
 RA Celniker S.E.;
 RT "Full length *Drosophila melanogaster* cDNA sequence."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003779; AAF57198.1; -
 DR EMBL; AF145677; AAD38652.1; -
 DR FlyBase; FBGN005632; faf.
 DR InterPro; IPR001394; UCH-2.
 DR Pfam; PF00442; UCH-1; 1.
 DR Pfam; PF00443; UCH-2; 1.
 DR PROSITE; PS00972; UCH_2_1; 1.
 DR PROSITE; PS00973; UCH_2_2; 1.
 DR PROSITE; PS0235; UCH_2_3; 1.
 SQ SEQUENCE 2778 AA; 311140 MW; FFB90438BA53A02B CRC64;
 Query Match 68.0%; Score 34; DB 5; Length 2778;
 Best Local Similarity 54.5%; Pred. No. 4.1e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDXS 11
 DB 1394 EVIIPDQDFS 1404
 I : I I I I I I I I
 I : I I I I I I I I
 RESULT 23
 Q8TX62 PRELIMINARY; PRT; 143 AA.
 AC Q8TX62;
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Uncharacterized conserved protein.
 GN MK0814.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN-AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;
 RT "The complete genome of hyperthermophile *Methanopyrus kandleri* AV19
 and monophyly of archaeal methanogens."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010372; AAM02027.1; -
 KW Complete proteome.
 SQ SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBEDD0B CRC64;

Query Match 66.0%; Score 33; DB 17; Length 143;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDX 10
 DB 75 EELVPGAGY 84
 I I I I I I I I
 I I I I I I I I
 RESULT 24
 Q92BC5 PRELIMINARY; PRT; 165 AA.
 AC Q92BC5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein lin1625.
 GN LIN1625.
 OS *Listeria innocua*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL; AL596169; CAC96856.1; -
 DR Listerist; LIN01625; -
 DR InterPro; IPR000866; AhpC-TSA.
 DR InterPro; IPR002065; Tpx.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; Tpx; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 18162 MW; 77705B7CD8BC8F4D CRC64;

Query Match 66.0%; Score 33; DB 16; Length 165;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXDX 10
 DB 144 EVVPEGSDH 152
 I I I I I I I I
 I I I I I I I I
 RESULT 25
 Q8Y6U8 PRELIMINARY; PRT; 165 AA.
 AC Q8Y6U8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein lmo1583.
 GN LMO1583.
 OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN-EGD-E / SEROVAR 1/2A;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL: AL591979; CAC99661.1; -.
 DR Listlist; LMO01583; -.
 DR InterPro; IPR000866; AhpC-TSA.
 DR InterPro; IPR002065; TPX.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 18133 MW; 77705B7CC46D424D CRC64;

Query Match . 66.0%; Score 33; DB 16; Length 165;
 Best Local Similarity 66.7%; Pred. No. 28;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10

|||||

Db 144 EVVPEGSDH 152

Search completed: June 10, 2003, 13:46:38
 Job time : 26.7857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds
(without alignments)
46.744 Million cell updates/sec

Title: US-09-909-164-46
Sequence: 1 EFVVPXDXYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : A_Geneseq_101002 : *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT : *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT : *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT : *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT : *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT : *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT : *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT : *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT : *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT : *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT : *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT : *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT : *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT : *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT : *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT : *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT : *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT : *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT : *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT : *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT : *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT : *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT : *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	94.0	11	23	Hepatitis C virus
2	47	94.0	11	23	Hepatitis C virus
3	47	94.0	11	23	Hepatitis C virus
4	47	94.0	11	23	Hepatitis C virus
5	47	94.0	11	23	Hepatitis C virus
6	46	92.0	11	23	Hepatitis C virus
7	46	92.0	11	23	Hepatitis C virus
8	46	92.0	11	23	Hepatitis C virus
9	46	92.0	11	23	Hepatitis C virus
10	46	92.0	11	23	Hepatitis C virus

11	46	92.0	11	23	ABB80538	Hepatitis C virus
12	46	92.0	11	23	ABB80542	Hepatitis C virus
13	46	92.0	11	23	ABB80543	Hepatitis C virus
14	46	92.0	11	23	ABB80561	Hepatitis C virus
15	46	92.0	11	23	ABB80562	Hepatitis C virus
16	41	82.0	11	23	ABB80544	Hepatitis C virus
17	41	82.0	11	23	ABB80545	Hepatitis C virus
18	41	82.0	11	23	ABB80549	Hepatitis C virus
19	41	82.0	11	23	ABB80552	Hepatitis C virus
20	41	82.0	11	23	ABB80553	Hepatitis C virus
21	40	80.0	11	23	ABB80521	Hepatitis C virus
22	40	80.0	11	23	ABB80522	Hepatitis C virus
23	40	80.0	11	23	ABB80525	Hepatitis C virus
24	40	80.0	11	23	ABB80526	Hepatitis C virus
25	40	80.0	11	23	ABB80530	Hepatitis C virus
26	40	80.0	11	23	ABB80535	Hepatitis C virus
27	40	80.0	11	23	ABB80536	Hepatitis C virus
28	40	80.0	11	23	ABB80539	Hepatitis C virus
29	40	80.0	11	23	ABB80540	Hepatitis C virus
30	40	80.0	11	23	ABB80546	Hepatitis C virus
31	40	80.0	11	23	ABB80550	Hepatitis C virus
32	40	80.0	11	23	ABB80554	Hepatitis C virus
33	40	80.0	11	23	ABB80555	Hepatitis C virus
34	40	80.0	11	23	ABB80559	Hepatitis C virus
35	40	80.0	11	23	ABB80563	Hepatitis C virus
36	40	80.0	11	23	ABB80564	Hepatitis C virus
37	40	80.0	11	23	ABB80565	Hepatitis C virus
38	40	80.0	11	23	ABB80566	Hepatitis C virus
39	40	80.0	11	23	ABB80567	Hepatitis C virus
40	40	80.0	11	23	ABB80568	Hepatitis C virus
41	39	78.0	11	23	ABB80523	Hepatitis C virus
42	39	78.0	11	23	ABB80527	Hepatitis C virus
43	39	78.0	11	23	ABB80531	Hepatitis C virus
44	39	78.0	11	23	ABB80532	Hepatitis C virus
45	39	78.0	11	23	ABB80537	Hepatitis C virus
46	39	78.0	11	23	ABB80541	Hepatitis C virus
47	39	78.0	11	23	ABB80558	Hepatitis C virus
48	39	78.0	11	23	ABB80560	Hepatitis C virus
49	35	70.0	143	22	ABB60256	Drosophila melanog
50	35	70.0	150	21	AAG28379	Arabidopsis thalia
51	35	70.0	299	22	ABG00168	Novel human thallo
52	35	70.0	587	21	RAY74287	Neisseria meningit
53	35	70.0	692	22	ABB64835	Drosophila melanog
54	34	68.0	150	22	AAH87623	Bovine mammary tis
55	34	68.0	175	21	AAG10069	Arabidopsis thalia
56	34	68.0	283	22	AAG11127	Novel central nerv
57	34	68.0	576	22	AAU87272	Novel human diagno
58	34	68.0	1022	22	ABG03621	Novel human diagno
59	34	68.0	1022	22	ABG05826	Novel human diagno
60	34	68.0	1022	22	ABG08173	Drosophila melanog
61	34	68.0	2778	22	ABB58683	Eubacterial DNA po
62	33	66.0	25	23	ABG62372	MAB L243 VH region
63	33	66.0	140	16	AAH64232	Humanized antibody
64	33	66.0	140	16	AAH64235	CDR-grafted L243-g
65	33	66.0	140	16	AAH64265	MHC-II MAB L243 he
66	33	66.0	140	16	AAH64257	Listeria monocytog
67	33	66.0	165	23	ABB48059	Group B Streptococ
68	33	66.0	222	22	AAU03629	Streptococcus poly
69	33	66.0	222	23	ABP26468	Murine JNK3 bindin
70	33	66.0	244	21	AAH12881	Murine JNK3 bindin
71	33	66.0	484	21	AAH12882	Streptococcus poly
72	33	66.0	563	23	ABP27069	Staphylococcus epi
73	32	64.0	139	23	ABP40279	Propionibacterium
74	32	64.0	150	22	AAU43140	Drosophila melanog
75	32	64.0	275	22	ABB63063	Drosophila melanog

ALIGNMENTS

RESULT 1
ID ABB80547 standard; peptide; 11 AA.

```

XX AC ABB80547;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #27.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX PX WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 94.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0081;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDXYS 11
Db 1 EEVVPXGTDYS 11
RESULT 2
ABB80548
ID ABB80548 standard; peptide; 11 AA.
XX AC ABB80548;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #28.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"

```

```

OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX PX WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 94.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0081;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDXYS 11
Db 1 EEVVPXGTDYS 11
RESULT 3
ABB80551
ID ABB80551 standard; peptide; 11 AA.
XX AC ABB80551;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX OS Synthetic.
XX FH Key
XX FT Modified-site 1 Location/Qualifiers
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"

```

FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
PI
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 94.0%; Score 47; DB 23; Length 11;
Best Local Similarity .90.9%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGSDYS 11
RESULT 4
ABB80556
ID ABB80556 standard; peptide; 11 AA.
XX
AC ABB80556;
XX
XX 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
PI
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 94.0%; Score 47; DB 23; Length 11;
Best Local Similarity .90.9%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGSDYS 11
RESULT 5
ABB80557
ID ABB80557 standard; peptide; 11 AA.
XX
AC ABB80557;
XX
XX 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX

XX
PF 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
PI
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 94.0%; Score 47; DB 23; Length 11;
Best Local Similarity .90.9%; Pred. No. 0.0081;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGSDYS 11
RESULT 5
ABB80557
ID ABB80557 standard; peptide; 11 AA.
XX
AC ABB80557;
XX
XX 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
PR (CORV-) CORVAS INT INC.
XX
XX

XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease -
XX CC
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX
XX Query Match 94.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0081;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EEVVPXGXDYS 11
XX ||||||| |||
XX Db 1 EEVVPXGSDYS 11
XX
XX RESULT 6
XX ABB80524
XX ID ABB80524 standard; peptide; 11 AA.
XX XX
XX AC ABB80524;
XX CC
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX FT
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX PS WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease -

XX PS Claim 17; Page 64; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX
XX Query Match 92.0%; Score 46; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.013;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EEVVPXGXDYS 11
XX ||||||| |||
XX Db 1 EEVVPXGMDYS 11
XX
XX RESULT 7
XX ABB80528
XX ID ABB80528 standard; peptide; 11 AA.
XX XX
XX AC ABB80528;
XX CC
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT Misc-difference 8 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX FT
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX PS WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease -
XX CC
XX PS Claim 17; Page 64; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.013;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDYS 11

Db 1 EEVVPXGMDYS 11

RESULT 8

ABB80529

ID ABB80529 standard; peptide; 11 AA.

XX AC ABB80529;

DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.

XX OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"

FT Misc-difference 8

FT Misc-difference 9 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.013;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDYS 11

Db 1 EEVVPXGMDYS 11

RESULT 9

ABB80533

ID ABB80533 standard; peptide; 11 AA.

XX AC ABB80533;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #13.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.

XX OS Synthetic.

PH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 92.0%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.013;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDYS 11

Db 1 EEVVPXGMDYS 11

RESULT 10

ABB80534

ID ABB80534 standard; peptide; 11 AA.
AC ABB80534;
XX
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #14.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 6
FT /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
XX
XX Misc-difference 9
FT /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
PD 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
PR
XX (CORV-) CORVAS INT INC.
PA
XX Lim-wilby M, Levy OE, Brunck TK;
PI WPI; 2002-361643/39.
DR
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 64; 69pp; English.
PS
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred No. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
Db 1 EEVVPXGGDYS 11
RESULT 11
ABB80538
ID ABB80538 standard; peptide; 11 AA.
XX
XX ABB80538;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
DE
XX

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 6
FT /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
XX
XX Misc-difference 9
FT /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO200208251-A2.
XX
PD 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
PR
XX (CORV-) CORVAS INT INC.
PA
XX Lim-wilby M, Levy OE, Brunck TK;
PI WPI; 2002-361643/39.
DR
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 64; 69pp; English.
PS
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
Db 1 EEVVPXGGDYS 11
RESULT 12
ABB80542
ID ABB80542 standard; peptide; 11 AA.
XX
XX ABB80542;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
DE
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl"
FT

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT PD
FT Modified-site 11 /note= "C-terminal amide"
FT
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease
FT
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXQDYS 11
Db 1 EEVVPXGXQDYS 11
RESULT 13
ABB80543
ID ABB80543 standard; peptide; 11 AA.
XX
XX ABB80543;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #23.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11

FT /note= "C-terminal amide"
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease
FT
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXQDYS 11
Db 1 EEVVPXGXQDYS 11
RESULT 14
ABB80561
ID ABB80561 standard; peptide; 11 AA.
XX
XX ABB80561;
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT
FT Modified-site 8 /note= "Oxymethionine"
FT Modified-site 11 /note= "C-terminal amide"
FT
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX

PF 19-JUL-2001; 2001WO-US23169.
PR 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. NO. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGMDYS 11
RESULT 15
ABB80562
ID ABB80562 standard; peptide; 11 AA.
XX ABB80562;
XX 08-OCT-2002 (first entry)
DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.
DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
KW Synthetic.
OS Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6
FT Modified-site 6
FT /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 8 /note= "Oxymethionine"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT WO200208251-A2.
PN 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX

PA (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ Query Match 92.0%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. NO. 0.013;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGXDYS 11
DB 1 EEVVPXGMDYS 11
RESULT 16
ABB80544
ID ABB80544 standard; peptide; 11 AA.
XX ABB80544;
XX 08-OCT-2002 (first entry)
DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.
DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
KW Synthetic.
OS Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6
FT Modified-site 6
FT /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Modified-site 11 /note= "C-terminal amide"
FT WO200208251-A2.
PN 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX

PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXYS 11
|||||||
Db 1 EEVVPXGTSYS 11

RESULT 17
ABB80545
ID ABB80545 standard; peptide; 11 AA.

XX ABB80545;
CC
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #25.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11
FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXYS 11
|||||||
Db 1 EEVVPXGTSYS 11

RESULT 18
ABB80549
ID ABB80549 standard; peptide; 11 AA.

XX ABB80549;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #29.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.

XX Synthetic.

XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 82.0%; Score 41; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGSSYS 11

RESULT 19

ABB80552

ID ABB80552 standard; peptide; 11 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGSSYS 11

RESULT 20

ABB80553

ID ABB80553 standard; peptide; 11 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGSSYS 11

RESULT 21

ABB80521

ID ABB80521 standard; peptide; 11 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGSSYS 11

RESULT 22

ABB80521

ID ABB80521 standard; peptide; 11 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGSSYS 11

RESULT 23

ABB80521

ID ABB80521 standard; peptide; 11 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 XX WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 PD 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 FT Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 XX Query Match 80.0%; Score 40; DB 23; Length 11;
 XX Best Local Similarity 81.8%; Pred. No. 0.2;
 XX Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db |||||||
 1 EEVVPXGMSYS 11
 RESULT 22
 ABB80522
 ID ABB80522 standard; peptide; 11 AA.
 XX ABB80522;
 AC
 XX 08-OCT-2002 (first entry)
 DT
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 PN 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 PD 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 FT Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 XX Query Match 80.0%; Score 40; DB 23; Length 11;
 XX Best Local Similarity 81.8%; Pred. No. 0.2;
 XX Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db |||||||
 1 EEVVPXGMSYS 11
 RESULT 23
 ABB80525
 ID ABB80525 standard; peptide; 11 AA.
 XX ABB80525;
 AC
 XX 08-OCT-2002 (first entry)
 DT
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "D-form residue"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.

XX 31-JAN-2002.
 PD 19-JUL-2001; 2001WO-US23169.
 PF 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 PI WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX Claim 17; Page 64; 69pp; English.
 PS The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 80.0%; Score 40; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGMSYS 11
 RESULT 24
 ABB80526
 ID ABB80526 standard; peptide; 11 AA.
 AC ABB80526;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 PD 19-JUL-2001; 2001WO-US23169.
 PF 21-JUL-2000; 2000US-220101P.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -

XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX Claim 17; Page 64; 69pp; English.
 PS The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 80.0%; Score 40; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 DB 1 EEVVPXGMSYS 11
 RESULT 25
 ABB80530
 ID ABB80530 standard; peptide; 11 AA.
 AC ABB80530;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #10.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT Modified-site 11 residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 PD 19-JUL-2001; 2001WO-US23169.
 PF 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -

XX PS Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
 Query Match 80.0%; Score 40; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EEVVPXGXDYS 11
 |||||
 Db 1 EEVVPXGGSYS 11

Search completed: June 10, 2003, 13:39:12
 Job time : 32.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds
(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-46
Perfect score: 50
Sequence: 1 EVVVPXGDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	66.0	140	4	US-08-569-147-76
2	33	66.0	140	4	US-08-569-147-82
3	32	64.0	139	4	US-09-134-001C-5124
4	31	62.0	622	2	US-08-459-146-2
5	31	62.0	622	2	US-08-459-065-2
6	31	62.0	739	4	US-09-413-814-86
7	30	60.0	59	4	US-08-963-851-14
8	30	60.0	121	4	US-09-152-060-68
9	30	60.0	121	4	US-09-152-060-85
10	30	60.0	122	2	US-08-879-995A-1
11	30	60.0	122	3	US-09-215-096-1
12	30	60.0	231	3	US-08-526-842B-20
13	30	60.0	240	3	US-08-926-842B-21
14	30	60.0	241	3	US-08-834-776A-2
15	30	60.0	341	4	US-08-853-948B-4
16	30	60.0	348	4	US-08-853-948B-5
17	30	60.0	368	4	US-09-697-367-24
18	30	60.0	421	4	US-09-093-448-4
19	30	60.0	478	2	US-09-040-799-3
20	30	60.0	478	4	US-09-093-448-1
21	30	60.0	478	4	US-09-093-448-2
22	30	60.0	478	4	US-09-093-448-3
23	30	60.0	1068	2	US-08-429-054A-11
24	30	60.0	1068	2	US-08-718-777-7
25	30	60.0	1068	3	US-09-051-341-7
26	29	58.0	100	2	US-09-047-125-27
27	29	58.0	100	3	US-07-736-335E-27

28	29	58.0	381	4	US-09-134-001C-3003	Sequence 3003, Ap
29	29	58.0	493	4	US-09-411-628-10	Sequence 10, Appl
30	29	58.0	543	4	US-09-739-455-4	Sequence 4, Appl
31	29	58.0	543	4	US-09-739-455-14	Sequence 14, Appl
32	29	58.0	544	4	US-09-732-025-4	Sequence 4, Appl
33	29	58.0	756	4	US-09-085-199B-9	Sequence 9, Appl
34	29	58.0	914	4	US-09-085-199B-4	Sequence 4, Appl
35	29	58.0	1090	4	US-09-085-199B-5	Sequence 5, Appl
36	28	56.0	45	2	US-08-637-739B-236	Sequence 236, App
37	28	56.0	45	3	US-08-871-355A-236	Sequence 236, App
38	28	56.0	45	4	US-09-201-945-236	Sequence 236, App
39	28	56.0	68	2	US-08-637-759B-62	Sequence 62, Appl
40	28	56.0	68	3	US-08-871-355A-62	Sequence 62, Appl
41	28	56.0	68	4	US-09-201-945-62	Sequence 62, Appl
42	28	56.0	181	4	US-09-134-001C-3897	Sequence 3897, Ap
43	28	56.0	216	4	US-09-198-119C-13	Sequence 13, Appl
44	28	56.0	364	4	US-09-338-671-2	Sequence 2, Appl
45	28	56.0	396	1	US-08-769-309A-15	Sequence 15, Appl
46	28	56.0	396	3	US-08-994-570-15	Sequence 15, Appl
47	28	56.0	453	1	US-08-769-309A-14	Sequence 14, Appl
48	28	56.0	453	3	US-08-994-570-14	Sequence 14, Appl
49	28	56.0	463	4	US-09-134-001C-3973	Sequence 3973, Ap
50	28	56.0	529	4	US-09-240-639-4	Sequence 4, Appl
51	28	56.0	534	2	US-08-878-563A-1	Sequence 1, Appl
52	28	56.0	534	4	US-09-270-117-1	Sequence 1, Appl
53	28	56.0	651	1	US-08-769-309A-17	Sequence 17, Appl
54	28	56.0	651	3	US-08-994-570-17	Sequence 17, Appl
55	28	56.0	746	4	US-09-149-934-4	Sequence 4, Appl
56	28	56.0	787	1	US-08-574-763-2	Sequence 2, Appl
57	28	56.0	873	3	US-08-990-140-2	Sequence 2, Appl
58	28	56.0	873	4	US-09-546-238-2	Sequence 2, Appl
59	28	56.0	947	4	US-09-228-986-73	Sequence 73, Appl
60	28	56.0	1167	1	US-08-100-709-2	Sequence 2, Appl
61	28	56.0	1167	1	US-08-176-865-2	Sequence 2, Appl
62	28	56.0	1167	1	US-08-474-038-2	Sequence 2, Appl
63	28	56.0	1167	2	US-08-779-046-2	Sequence 2, Appl
64	28	56.0	1167	2	US-08-881-340-2	Sequence 2, Appl
65	28	56.0	1722	4	US-09-194-612A-1	Sequence 1, Appl
66	28	56.0	1780	1	US-08-769-309A-5	Sequence 5, Appl
67	28	56.0	1780	3	US-08-994-570-5	Sequence 5, Appl
68	27	54.0	24	4	US-09-227-357-403	Sequence 403, App
69	27	54.0	126	2	US-08-879-995A-3	Sequence 3, Appl
70	27	54.0	126	3	US-09-215-096-3	Sequence 3, Appl
71	27	54.0	159	2	US-08-844-086-4	Sequence 4, Appl
72	27	54.0	159	3	US-09-018-211-4	Sequence 4, Appl
73	27	54.0	168	3	US-08-483-534A-2	Sequence 2, Appl
74	27	54.0	179	2	US-08-934-959-3	Sequence 3, Appl
75	27	54.0	195	2	US-08-960-022-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-569-147-76
; Sequence 76, Application US/08569147
; Patent No. 6180377
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMANISED ANTIBODIES
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESS: No. 6180377ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,147
FILING DATE: 25-March-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0047
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-147-76

Query Match 66.0%; Score 33; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDY 10
||| |
Db 122 VVPTGFDY 129

RESULT 2
US-08-569-147-82
Sequence 82, Application US/08569147
Patent No. 6180377
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMANISED ANTIBODIES
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6180377ris, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,147
FILING DATE: 25-March-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0047
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-147-82

Query Match 66.0%; Score 33; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDY 10
||| |
Db 122 VVPTGFDY 129

RESULT 3
US-09-134-001C-5124
Sequence 5124, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5124
LENGTH: 139
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5124

Query Match 64.0%; Score 32; DB 4; Length 139;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
||| |
Db 32 IVPFGHDYN 40

RESULT 4
US-08-459-146-2
Sequence 2, Application US/08459146
Patent No. 5866405
GENERAL INFORMATION:
APPLICANT: Choi, Gil Ho
APPLICANT: Nuss, Donald Lee
TITLE OF INVENTION: Genetically Engineered Transmissible
TITLE OF INVENTION: Hypovirulence
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,146
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/832,117
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Roseman, Catherine R
REGISTRATION NUMBER: 34,240
REFERENCE/DOCKET NUMBER: 8589
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-6208
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 622 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Endothia parasitica (Cryptonectria)
ORGANISM: parasitica)
STRAIN: EP713
US-08-459-146-2

Query Match 62.0%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
Db 31 EEVVPAG 37

RESULT 5

US-08-459-065-2
Sequence 2, Application US/08459065
Patent No. 5882642

GENERAL INFORMATION:

APPLICANT: Choi, Gil Ho
TITLE OF INVENTION: Genetically Engineered Transmissible
TITLE OF INVENTION: Hypovirulence
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,065
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/832,117
FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Roseman, Catherine R
REGISTRATION NUMBER: 34,240
REFERENCE/DOCKET NUMBER: 8589
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-6208
TELEFAX: (201) 235-3500

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 622 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Endothia parasitica (Cryptonectria)
ORGANISM: parasitica)
STRAIN: EP713

US-08-459-065-2

Query Match 62.0%; Score 31; DB 2; Length 622;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0;

0; Gaps

0; Indels

0; Mismatches

0; Conservative

0; Length

0; Pred. No.

Qy 1 EEVVPXG 7
Db 31 EEVVPAG 37

RESULT 6

US-09-413-814-86
Sequence 86, Application US/09413814
Patent No. 6225064

GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolyketide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 86
LENGTH: 739
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-86

Query Match 62.0%; Score 31; DB 4; Length 739;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VPXGXDXYS 11
Db 663 IPLGGDYS 670

RESULT 7

US-08-963-851-14
Sequence 14, Application US/08963851
Patent No. 6300116

GENERAL INFORMATION:

APPLICANT: VAN DER OSTEN, CLAUS
APPLICANT: HALKIER, TORDEN
APPLICANT: ANDERSEN, CARSTEN
APPLICANT: BAUDITZ, PETER
APPLICANT: HANSEN, PETER KAMP
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
FILE REFERENCE: 4946,200-US
CURRENT APPLICATION NUMBER: US/08/963,851
CURRENT FILING DATE: 1997-11-04
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 59
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-08-963-851-14

Query Match 60.0%; Score 30; DB 4; Length 59;
Best Local Similarity 45.5%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDXYS 11

Db 38 EKHIPGLEYS 48
|: :| | :||

RESULT 8
US-09-152-060-68
; Sequence 68, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-152-060-68

Query Match 60.0%; Score 30; DB 4; Length 121;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

RESULT 9
US-09-152-060-85
; Sequence 85, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30

; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-152-060-85

Query Match 60.0%; Score 30; DB 4; Length 121;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
QY 1 EEVVPXG 7
| | | | |
Db 28 EEVVPXG 34

RESULT 10
US-08-879-995A-1
; Sequence 1, Application US/08879995A
; Patent No. 5985606
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/879,995A
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2109906
US-08-879-995A-1

Query Match 60.0%; Score 30; DB 2; Length 122;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
DB 28 EEVPPGG 34

RESULT 11
US-09-215-096-1
; Sequence 1, Application US/09215096
; Patent No. 6008194
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215.096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/879,995
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2109906
US-09-215-096-1

Query Match 60.0%; Score 30; DB 3; Length 122;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
| | | | |
DB 28 EEVPPGG 34

US-08-926-842B-20
; Sequence 20, Application US/08926842B
; Patent No. 6030807
; GENERAL INFORMATION:
; APPLICANT: Sa-No. 6030807ueira, Isabel
; APPLICANT: de Lencastre, Herminia
; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926.842B
; FILING DATE: 10-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-089 N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
US-08-926-842B-20

Query Match 60.0%; Score 30; DB 3; Length 231;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
: | | | |
DB 41 IKPSGVDYS 49

RESULT 13
US-08-926-842B-21
; Sequence 21, Application US/08926842B
; Patent No. 6030807
; GENERAL INFORMATION:
; APPLICANT: Sa-No. 6030807ueira, Isabel
; APPLICANT: de Lencastre, Herminia
; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```

; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-834-776A-2

Query Match          60.0%; Score 30; DB 3; Length 241;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 EEVVPXG 7
DB      199 EEIVPAG 205
      ||:|||
RESULT 15
US-08-853-948B-4
; Sequence 4, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Citrus unshiu
; FEATURE:
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asn, Asp,
; OTHER INFORMATION: Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe,
; OTHER INFORMATION: Pro, Ser, Thr, Trp, Tyr, or Val
US-08-853-948B-4

Query Match          60.0%; Score 30; DB 4; Length 341;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGXDXYS 11
DB      228 VIPPGMDFS 236
      |:| |:|
RESULT 16
US-08-853-948B-5
; Sequence 5, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Citrus unshiu
US-08-853-948B-5

Query Match          60.0%; Score 30; DB 4; Length 348;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGXDXYS 11
      |:| |:|

```

Db 234 VIPPGMDFS 242

RESULT 17

US-09-697-367-24

; Sequence 24, Application US/09697367

; Patent No. 6323015

; GENERAL INFORMATION:

; APPLICANT: Orozco Jr., Emil M.

; APPLICANT: Caimi, Perry G.

; APPLICANT: Weng, Zude

; APPLICANT: Tarczynski, Mitchell

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE

; FILE REFERENCE: Bb1166 US NA

; CURRENT APPLICATION NUMBER: US/09/697,367

; CURRENT FILING DATE: 2000-10-26

; PRIOR APPLICATION NUMBER: 60/084,529

; PRIOR FILING DATE: 1998-MAY-07

; PRIOR APPLICATION NUMBER: PCT/US99/09865

; PRIOR FILING DATE: 1999-MAY-06

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; LENGTH: 368

; TYPE: PRT

; ORGANISM: Zea mays

US-09-697-367-24

Query Match 60.0%; Score 30; DB 4; Length 368;

Best Local Similarity 55.6%; Pred. No. 1.9e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGXDXYS 11

1:111111

Db 217 VIPPGMDFS 225

RESULT 18

US-09-093-448-4

; Sequence 4, Application US/09093448A

; Patent No. 6207704

; GENERAL INFORMATION:

; APPLICANT: Liu, Jun O.

; APPLICANT: Griffith, Eric C.

; APPLICANT: Su, Zhuang

; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors

; TITLE OF INVENTION: and Uses Thereof

; FILE REFERENCE: 0492611-0346

; CURRENT APPLICATION NUMBER: US/09/093,448A

; CURRENT FILING DATE: 1998-06-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 421

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-093-448-4

Query Match 60.0%; Score 30; DB 4; Length 421;

Best Local Similarity 60.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10

:111111

Db 412 KEVSRGDDY 421

RESULT 19

US-09-040-799-3

; Sequence 3, Application US/09040799

; Patent No. 5885820

; GENERAL INFORMATION:

; APPLICANT: CHANG, YIE-HWA

; TITLE OF INVENTION: CLONE OF A NUCLEOTIDE SEQUENCE ENCODING

; TITLE OF INVENTION: A PROTEIN HAVING TWO FUNCTIONS

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HOWELL & HAERKAMP, L.C.

; STREET: 7733 FORSYTH BLVD., SUITE 1400

; CITY: ST. LOUIS

; STATE: MO

; COUNTRY: USA

; ZIP: 63105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/040,799

; FILING DATE: 18-MAR-1998

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: HOLLAND, DONALD R

; REGISTRATION NUMBER: 35,197

; REFERENCE/DOCKET NUMBER: 16153-4639

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 314-727-5188

; TELEFAX: 314-727-6092

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 478 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-040-799-3

Query Match 60.0%; Score 30; DB 2; Length 478;

Best Local Similarity 60.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10

:111111

Db 469 KEVSRGDDY 478

RESULT 20

US-09-093-448-1

; Sequence 1, Application US/09093448A

; Patent No. 6207704

; GENERAL INFORMATION:

; APPLICANT: Liu, Jun O.

; APPLICANT: Griffith, Eric C.

; APPLICANT: Su, Zhuang

; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors

; TITLE OF INVENTION: and Uses Thereof

; FILE REFERENCE: 0492611-0346

; CURRENT APPLICATION NUMBER: US/09/093,448A

; CURRENT FILING DATE: 1998-06-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 478

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-093-448-1

Query Match 60.0%; Score 30; DB 4; Length 478;

Best Local Similarity 60.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGXDX 10

:111111

Db 469 KEVSRGDDY 478

RESULT 21

US-09-093-448-2
; Sequence 2, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 478
; ORGANISM: Rattus norvegicus
US-09-093-448-2

Query Match 60.0%; Score 30; DB 4; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
; :||| | ||
Db 469 KEVVRGDDY 478

RESULT 22

US-09-093-448-3
; Sequence 3, Application US/09093448A
; Patent No. 6207704
; GENERAL INFORMATION:
; APPLICANT: Liu, Jun O.
; APPLICANT: Griffith, Eric C.
; APPLICANT: Su, Zhuang
; TITLE OF INVENTION: Type 2 Methionine Aminopeptidase (MetAP2) Inhibitors
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 0492611-0346
; CURRENT APPLICATION NUMBER: US/09/093,448A
; CURRENT FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-093-448-3

Query Match 60.0%; Score 30; DB 4; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
; :||| | ||
Db 469 KEVVRGDDY 478

RESULT 23

US-08-429-054A-11
; Sequence 11, Application US/08429054A
; Patent No. 5917126
; GENERAL INFORMATION:
; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,
; APPLICANT: JEAN; VOELKER, TONI; GERVAIS, MONICA
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),
; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE
; TITLE OF INVENTION: CDNA FOR MODIFYING SPS EXPRESSION IN PLANT CELLS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN AND MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,054A
; FILING DATE: 26-APR-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 842,337
; FILING DATE: 20-March-1992
; APPLICATION NUMBER: PCT/FR 91/00593
; FILING DATE: 18-July-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: French 90402094.9
; FILING DATE: 20-July-1990
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles A. Muserlian
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068
; TYPE: Amino acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
US-08-429-054A-11

Query Match 60.0%; Score 30; DB 2; Length 1068;
Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
; :| | | | :|
Db 435 VIPPGMDFS 443

RESULT 24

US-08-718-777-7
; Sequence 7, Application US/08718777
; Patent No. 5981852
; GENERAL INFORMATION:
; APPLICANT: Van Assche, C.
; APPLICANT: Lando, D.
; APPLICANT: Bruneau, J. M.
; APPLICANT: Voelker, T.
; APPLICANT: Gervais, M.
; TITLE OF INVENTION: MODIFICATION OF SUCROSE
; TITLE OF INVENTION: PHOSPHATE
; TITLE OF INVENTION: SYNTHASE IN PLANTS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Barbara Rae-Venter
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,777
; FILING DATE: NOT YET ASSIGNED
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,471
; FILING DATE: 27-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.072.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400
; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-718-777-7

Query Match 60.0%; Score 30; DB 2; Length 1068;

Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
|:| | |:
Db 435 VIPPGMDFS 443

RESULT 25

US-09-051-341-7
; Sequence 7, Application US/09051341
; Patent No. 6124528
; GENERAL INFORMATION:
; APPLICANT: Shewmaker, C. K.
; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,341
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/17351
; FILING DATE: 25-OCT-1996
; APPLICATION NUMBER: US 08/549,016
; FILING DATE: 27-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,200
; FILING DATE: 12-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter, Ph.D.,
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.110.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400

; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-051-341-7

Query Match 60.0%; Score 30; DB 3; Length 1068;
Best Local Similarity 55.6%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
|:| | |:
Db 435 VIPPGMDFS 443

Search completed: June 10, 2003, 13:51:38
Job time : 9.64286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds
(without alignments)
75.710 Million cell updates/sec

Title: US-09-909-164-46
Perfect score: 50
Sequence: 1 BEVVPXGXDYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	68.0	283	9	US-09-738-626-4881
2	33	66.0	440	9	US-09-813-408-27
3	32	64.0	3472	9	US-10-027-806-4
4	32	64.0	3472	9	US-10-034-623-4
5	32	64.0	3472	9	US-10-027-801-4
6	31	62.0	299	10	US-09-815-242-10697
7	31	62.0	653	9	US-09-820-843A-26
8	31	62.0	846	10	US-09-815-242-13904
9	30	60.0	7	9	US-09-909-062-1
10	30	60.0	7	9	US-09-909-062-9
11	30	60.0	7	9	US-09-909-062-130
12	30	60.0	11	9	US-09-943-123-22
13	30	60.0	59	10	US-09-948-080-14
14	30	60.0	79	9	US-09-764-891-4601
15	30	60.0	121	9	US-09-852-797-68
16	30	60.0	121	9	US-09-852-797-85
17	30	60.0	121	10	US-09-853-161-68
18	30	60.0	121	10	US-09-853-161-85
19	30	60.0	121	10	US-09-852-659A-68

20	30	60.0	121	10	US-09-852-659A-85	Sequence 85, Appl
21	30	60.0	135	9	US-09-992-598-359	Sequence 359, App
22	30	60.0	135	9	US-09-989-293A-359	Sequence 359, App
23	30	60.0	135	9	US-09-989-735-359	Sequence 359, App
24	30	60.0	135	9	US-09-990-444-359	Sequence 359, App
25	30	60.0	135	9	US-09-989-730-359	Sequence 359, App
26	30	60.0	135	9	US-09-990-436-359	Sequence 359, App
27	30	60.0	135	9	US-09-991-181-359	Sequence 359, App
28	30	60.0	135	9	US-09-993-687-359	Sequence 359, App
29	30	60.0	135	9	US-09-989-734-359	Sequence 359, App
30	30	60.0	135	9	US-09-997-653-359	Sequence 359, App
31	30	60.0	135	9	US-10-174-590-444	Sequence 444, App
32	30	60.0	135	9	US-10-176-758-444	Sequence 444, App
33	30	60.0	135	9	US-10-175-737-444	Sequence 444, App
34	30	60.0	135	9	US-09-993-667-359	Sequence 359, App
35	30	60.0	135	9	US-10-173-706-444	Sequence 444, App
36	30	60.0	135	9	US-10-175-738-444	Sequence 444, App
37	30	60.0	135	9	US-10-175-752-444	Sequence 444, App
38	30	60.0	135	9	US-10-176-482-444	Sequence 444, App
39	30	60.0	135	9	US-10-176-757-444	Sequence 444, App
40	30	60.0	135	9	US-10-176-913-444	Sequence 444, App
41	30	60.0	135	9	US-10-180-552-444	Sequence 444, App
42	30	60.0	135	9	US-10-180-557-444	Sequence 444, App
43	30	60.0	135	9	US-09-990-438-359	Sequence 359, App
44	30	60.0	135	9	US-09-990-562-359	Sequence 359, App
45	30	60.0	135	9	US-09-997-428-359	Sequence 359, App
46	30	60.0	135	9	US-09-997-666-359	Sequence 359, App
47	30	60.0	135	9	US-10-173-700-444	Sequence 444, App
48	30	60.0	135	9	US-10-174-572-444	Sequence 444, App
49	30	60.0	135	9	US-10-174-579-444	Sequence 444, App
50	30	60.0	135	9	US-10-174-582-444	Sequence 444, App
51	30	60.0	135	9	US-10-174-588-444	Sequence 444, App
52	30	60.0	135	9	US-10-175-739-444	Sequence 444, App
53	30	60.0	135	9	US-10-175-740-444	Sequence 444, App
54	30	60.0	135	9	US-10-175-743-444	Sequence 444, App
55	30	60.0	135	9	US-10-176-488-444	Sequence 444, App
56	30	60.0	135	9	US-10-176-492-444	Sequence 444, App
57	30	60.0	135	9	US-10-176-747-444	Sequence 444, App
58	30	60.0	135	9	US-10-176-750-444	Sequence 444, App
59	30	60.0	135	9	US-10-176-985-444	Sequence 444, App
60	30	60.0	135	9	US-10-176-987-444	Sequence 444, App
61	30	60.0	135	9	US-10-176-991-444	Sequence 444, App
62	30	60.0	135	9	US-10-176-992-444	Sequence 444, App
63	30	60.0	135	9	US-10-176-993-444	Sequence 444, App
64	30	60.0	135	9	US-10-184-658-444	Sequence 444, App
65	30	60.0	135	9	US-10-227-884-108	Sequence 108, App
66	30	60.0	135	9	US-09-990-711-359	Sequence 359, App
67	30	60.0	135	9	US-10-173-695-444	Sequence 444, App
68	30	60.0	135	9	US-10-173-697-444	Sequence 444, App
69	30	60.0	135	9	US-10-173-705-444	Sequence 444, App
70	30	60.0	135	9	US-10-174-576-444	Sequence 444, App
71	30	60.0	135	9	US-10-174-585-444	Sequence 444, App
72	30	60.0	135	9	US-10-174-586-444	Sequence 444, App
73	30	60.0	135	9	US-10-175-747-444	Sequence 444, App
74	30	60.0	135	9	US-10-176-481-444	Sequence 444, App
75	30	60.0	135	9	US-10-176-485-444	Sequence 444, App

ALIGNMENTS

RESULT 1
US-09-738-626-4881
; Sequence 4881, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIKOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO

; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4881
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4881

Query Match 68.0%; Score 34; DB 9; Length 283;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGXDYS 11
||| |||
Db 56 VPAGADYS 63

RESULT 2

US-09-813-408-27
; Sequence 27, Application US/09813408
; Publication No. US20030049619A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Maris, Barry
; TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides
; TITLE OF INVENTION: Of Polynucleotides
; FILE REFERENCE: HER0041
; CURRENT APPLICATION NUMBER: US/09/813,408
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match 66.0%; Score 33; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10
||| |||
Db 120 EVLPWGVY 128

RESULT 3

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/027,806
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29

; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 64.0%; Score 32; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
||| |||
Db 2294 EDVIPRGISFS 2304

RESULT 4

US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 64.0%; Score 32; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
||| |||
Db 2294 EDVIPRGISFS 2304

RESULT 5

US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/027,801
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match 64.0%; Score 32; DB 9; Length 3472;

Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDXS 11
Db 2294 EDVIPRGISFS 2304

RESULT 6

US-09-815-242-10697
; Sequence 10697, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10697
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Enterococcus faecalis

US-09-815-242-10697

Query Match 62.0%; Score 31; DB 10; Length 299;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDX 10
Db 218 EQITPTGIEY 227

RESULT 7

US-09-820-843A-26

; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1

; GENERAL INFORMATION:

; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES

; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 653
; TYPE: PRT

; ORGANISM: Vibrio cholerae
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Iron(III) ABC transporter, permease protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: gi|9654609
US-09-820-843A-26

Query Match 62.0%; Score 31; DB 9; Length 653;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
Db 300 EEVVPXG 306

RESULT 8

US-09-815-242-13904
; Sequence 13904, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13904
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Salmonella typhi

US-09-815-242-13904

Query Match 62.0%; Score 31; DB 10; Length 846;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDX 10
Db 762 EDVDMGVY 771

RESULT 9

US-09-909-062-1

; Sequence 1, Application US/09909062
; Publication No. US20030036501A1
; GENERAL INFORMATION:
; APPLICANT: Saksena, Anil K
; APPLICANT: Girijavallabhan, Viyyor M

APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-1

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
DB 1 EEVVPXG 7

RESULT 10
US-09-909-062-9
Sequence 9, Application US/09909062
Publication No. US20030036501A1
GENERAL INFORMATION:
APPLICANT: Saksena, Anil K
APPLICANT: Girijavaliabhan, Viyyor M
APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 7
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-9

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
DB 1 EEVVPXG 7

RESULT 11
US-09-909-062-130
Sequence 130, Application US/09909062
Publication No. US20030036501A1
GENERAL INFORMATION:
APPLICANT: Saksena, Anil K
APPLICANT: Girijavaliabhan, Viyyor M
APPLICANT: Lovey, Raymond G
APPLICANT: Jao, Edwin
APPLICANT: Bennett, Frank
APPLICANT: McCormick, Jinping L
APPLICANT: Pike, Russell E
APPLICANT: Bogen, Stephane L
APPLICANT: Liu, Yi-Tsung
APPLICANT: Arasappan, Ashok
APPLICANT: Pinto, Patrick A
APPLICANT: Njoroge, F George
APPLICANT: Ganguly, Ashit
TITLE OF INVENTION: NOVEL PEPTIDES AS NS3-SERINE PROTEASE INHIBITORS OF HEPATITIS C
FILE REFERENCE: IN01157K-US
CURRENT APPLICATION NUMBER: US/09/909,062
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/220,109
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn version 3.1
SEQ ID NO 130
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic peptide
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (7)..(7)
OTHER INFORMATION: AMIDATION
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (6)..(6)
OTHER INFORMATION: norvaline-C(-O)
US-09-909-062-130

Query Match 60.0%; Score 30; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXG 7

Db 1 EEVVPXG 7
|||||

RESULT 12

US-09-943-123-22
; Sequence 22, Application US/09943123
; Publication No. US20020182701A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Y-H
; APPLICANT: VETRO, J.A.
; APPLICANT: MICKA, W.S.
; TITLE OF INVENTION: Dominant Negative Variants of Methionine Aminopeptidase
; FILE OF INVENTION: 2 ("MetAP2") and Clinical Uses Therefor
; FILE REFERENCE: 16153-8007
; CURRENT APPLICATION NUMBER: US/09/943,123
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-943-123-22

Query Match 60.0%; Score 30; DB 9; Length 11;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
|:|:| | | | |
Db 2 KEVSKGDDY 11

RESULT 13

US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102703A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match 60.0%; Score 30; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
|: | | | | |
Db 38 EKHIPGLEYS 48

RESULT 14

US-09-764-891-4601
; Sequence 4601, Application US/09764891

; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4601
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-891-4601

Query Match 60.0%; Score 30; DB 9; Length 79;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EEVVPXGDYS 11
|:|:| | | | |
Db 41 EIIPIGSDYS 50

RESULT 15

US-09-852-797-68
; Sequence 68, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-797-68

Query Match 60.0%; Score 30; DB 9; Length 121;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXG 7
|||||

Db 28 EEVPPG 34

RESULT 16

US-09-852-797-85
; Sequence 85, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-797-85

Query Match 60.0%; Score 30; DB 9; Length 121;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVPPG 7

Db 28 EEVPPG 34

RESULT 17

US-09-853-161-68
; Sequence 68, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-161-68

Query Match 60.0%; Score 30; DB 10; Length 121;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVPPG 7

Db 28 EEVPPG 34

RESULT 18

US-09-853-161-85
; Sequence 85, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 121
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE

; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (89)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-853-161-85

Query Match 60.0%; Score 30; DB 10; Length 121;

Best Local Similarity 85.7%; Pred. No. 82;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7

Db 28 EEVPPGG 34

RESULT 19

US-09-852-659A-68

; Sequence 68, Application US/09852659A

; Patent No. US20020077287A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P4

; CURRENT APPLICATION NUMBER: US/09/852,659A

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 68

; LENGTH: 121

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-852-659A-68

Query Match 60.0%; Score 30; DB 10; Length 121;

Best Local Similarity 85.7%; Pred. No. 82;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7

Db 28 EEVPPGG 34

RESULT 20

US-09-852-659A-85

; Sequence 85, Application US/09852659A

; Patent No. US20020077287A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: P2003P4

; CURRENT APPLICATION NUMBER: US/09/852,659A

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 85

; LENGTH: 121

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (67)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (89)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-659A-85

Query Match 60.0%; Score 30; DB 10; Length 121;

Best Local Similarity 85.7%; Pred. No. 82;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7

Db 28 EEVPPGG 34

RESULT 21

US-09-992-598-359

; Sequence 359, Application US/09992598

; Patent No. US20020160384A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tamas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730PIC20
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217

; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088861
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/088876
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: 60/089105
 ; PRIOR FILING DATE: 1998-06-12
 ; PRIOR APPLICATION NUMBER: 60/089440
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089512
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089514
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089532
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089538
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089598
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089599
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089600
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089653
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089907
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;

Best Local Similarity 85.7%; Pred. No. 93;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPKG 7

|||||

Db 28 EEVVPKG 34

RESULT 22

US-09-989-293A-359

; Sequence 359, Application US/09989293A

; Patent No. US2002017716A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC66
; CURRENT APPLICATION NUMBER: US/09/989,293A
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVFXG 7
DB 28 EEVFXG 34

RESULT 23
US-09-989-735-359
; Sequence 359, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC61
CURRENT APPLICATION NUMBER: US/09/989,735
CURRENT FILING DATE: 2001-11-19
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;

Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXG 7
 |||||
Db 28 EEVVPXG 34

RESULT 24

US-09-990-444-359
; Sequence 359, Application US/09990444
; Publication No. US20020193300A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC19
; CURRENT APPLICATION NUMBER: US/09/990.444

;; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 60.08; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXG 7
|||||
Db 28 EEVVPXG 34

RESULT 25

; Sequence 359, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC69
; CURRENT APPLICATION NUMBER: US/09989,730
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12

;	PRIOR FILING DATE:	1998-06-16
;	PRIOR APPLICATION NUMBER:	60/089532
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089538
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089598
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089599
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089600
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089653
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089801
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089907
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089908
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089947
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089948
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089952
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/090246
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090252
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090254
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090349
;	PRIOR FILING DATE:	1998-06-23
;	PRIOR APPLICATION NUMBER:	60/090355
;	PRIOR FILING DATE:	1998-06-23
;	PRIOR APPLICATION NUMBER:	60/090429
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090431
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090435
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090444
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090445
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090472
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090535
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090540
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090542
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090690
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090694
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090676
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090678
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090690
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090862
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/090863
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/091360
;	PRIOR FILING DATE:	1998-07-01

; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 60.0%; Score 30; DB 9; Length 135;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXG 7
Db 28 EEVPCG 34

Search completed: June 10, 2003, 14:35:46
Job time : 15.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds

(without alignments)

94.297 Million cell updates/sec

Title: US-09-909-164-46

Perfect score: 50

Sequence: 1 EEVWPXGDYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	72.0	363	D69551	conserved hypotet
2	35	70.0	290	D98182	6-O-methylguanine-D
3	35	70.0	290	AG3104	6-O-methylguanine-D
4	35	70.0	587	F81138	succinate dehydrog
5	35	70.0	906	T48898	disease resistance
6	35	70.0	908	T48899	disease resistance
7	34	68.0	102	A42452	V1 protein - tobac
8	34	68.0	156	S54619	hypothetical prote
9	34	68.0	247	A96001	conserved hypotet
10	34	68.0	257	A96546	unknown protein [i
11	34	68.0	394	F82491	ferrisiderophore r
12	34	68.0	433	H87660	peptidoglycan-bind
13	34	68.0	2747	B49132	fat facets (fat) s
14	33	66.0	124	VKLJ51	trans-regulatory s
15	33	66.0	165	AG1272	thiol peroxidases
16	33	66.0	165	AH1635	thiol peroxidases
17	33	66.0	196	AD0454	conserved hypotet
18	33	66.0	225	S57810	hypothetical prote
19	33	66.0	327	S40753	hypothetical prote
20	33	66.0	421	DEXTCM	acyl-CoA dehydroge
21	33	66.0	440	H72784	probable alkaline
22	33	66.0	1028	AF3286	ATP-dependent DNA
23	33	66.0	1088	D82246	probable chitinase
24	33	66.0	1150	T20173	hypothetical prote
25	32	64.0	99	S00210	plastocyanin b - L
26	32	64.0	155	S38255	plastocyanin b pre
27	32	64.0	168	S58208	conserved hypotet
28	32	64.0	196	AI0931	hypothetical prote
29	32	64.0	301	F89957	hypothetical prote

ALIGNMENTS

RESULT 1

D69551 conserved hypothetical protein AF2411 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: D69551

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.;

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Syk

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing a

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69551

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-363 <KLE>

A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AAB91255.1; PID:g

Query Match 72.0%; Score 36; DB 2; Length 363;

Best Local Similarity 54.5%; Pred. No. 9;

hypothetical prote
hypothetical prote
probable hexosyltr
L-lactate dehydrog
hypothetical prote
tolB protein - Hae
ABC transporter AT
oligopeptidase [im
protein B0212.3 [1
probable membrane
hypothetical 367K
projectin - fruit
S-adenosylmethioni
conserved hypotet
hypothetical prote
hypothetical prote
acyl-CoA dehydroge
hypothetical prote
probable succinate
hypothetical prote
iron(III) ABC tran
penicillin-binding
penicillin-binding
hemolysin - Synec
hypothetical prote
capsid protein - f
Ig heavy chain V r
plastocyanin - car
partial transposas
hypothetical prote
hypothetical prote
transcription fact
fimbrial chain [im
hypothetical prote
hypothetical prote
general stress pro
L-ribulose-phospha
L-ribulose-5-phosp
L-ribulose-5-phosp
L-ribulose-5-phosp
3-oxoacid CoA-tran
3-oxoadipate coA-t
L-ribulose-phospha

32 64.0 307 2 F84330
31 64.0 314 2 AH1912
32 64.0 357 1 G69290
33 64.0 366 2 G69350
34 64.0 425 2 T24111
35 64.0 427 2 F64064
36 64.0 565 2 E86665
37 64.0 632 2 H84350
38 64.0 672 2 G88651
39 64.0 1474 2 F69009
40 64.0 1472 2 T31308
41 64.0 6658 2 T13931
42 64.0 117 2 A69487
43 64.0 202 2 H97247
44 64.0 233 2 E72330
45 64.0 296 2 F72745
46 64.0 319 2 S03833
47 64.0 395 2 H84113
48 64.0 421 1 DEHUCM
49 64.0 496 2 S76296
50 64.0 587 2 D81881
51 64.0 622 2 S15009
52 64.0 630 2 F85074
53 64.0 653 2 D82352
54 64.0 840 2 AG0526
55 64.0 846 2 S57580
56 64.0 1741 2 S74910
57 64.0 13055 2 T16580
30.5 61.0 668 2 JQ2356
59 60.0 21 2 B49042
60 60.0 97 2 JW0011
61 60.0 97 2 A9427
62 60.0 128 2 A90471
63 60.0 165 2 D69493
64 60.0 175 2 S36749
65 60.0 180 2 AG0504
66 60.0 184 2 B86192
67 60.0 184 2 E90335
68 60.0 204 2 S66082
69 60.0 231 1 ISECP4
70 60.0 231 2 A90637
71 60.0 231 2 A85488
72 60.0 231 2 AB0515
73 60.0 232 2 C71908
74 60.0 232 2 C64606
75 60.0 248 1 ISEB4T

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 11
| : | | | : |
Db 120 ENIVPGIDFS 130

RESULT 2
D98182
O6-methylguanine-DNA methyltransferase PA2118 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98182
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, D.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens C58
A:Reference number: A97359; PMID:11743194
A:Accession: D98182
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK88982.1; PID:g15158766; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_818
A:Map position: linear chromosome

Query Match 70.0%; Score 35; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| : | | | : |
Db 9 EDITPIGSDY 18

RESULT 3
AG3104
6-O-methylguanine-DNA methyltransferase [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG3104
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, S.; Romero, P.; Zhang, S.; Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.E.;
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG3104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL45253.1; PID:g17742937; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: ada
A:Map position: linear chromosome

Query Match 70.0%; Score 35; DB 2; Length 290;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| : | | | : |
Db 9 EDITPIGSDY 18

RESULT 4
F81138
succinate dehydrogenase, flavoprotein chain NMB0950 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81138
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, R.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.; Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <TET>
A:Cross-references: GB:AE002446; GB:AE002098; NID:g7226185; PIDN:AAF41356.1; PID:g7226185
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0950
C:Superfamily: fumarate reductase flavoprotein; 3-oxosteroid 1-dehydrogenase homol

Query Match 70.0%; Score 35; DB 2; Length 587;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| : | | | : |
Db 366 EVVPPQGEDY 375

RESULT 5
T48898
disease resistance protein RPP8 [validated] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 21-Jul-2000
C:Accession: T48898
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Plant Cell 10, 1861-1874, 1998
A:Title: Intragenic recombination and diversifying selection contribute to the evolution of the Arabidopsis thaliana RPP8 gene
A:Reference number: Z24999; MUID:99030193; PMID:9811794
A:Accession: T48898
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-906 <MCD>
A:Cross-references: EMBL:AF089710; NID:g3928861; PIDN:AAC83165.1; PID:g3928862
A:Experimental source: Landsberg erecta
C:Genetics:
A:Gene: RPP8
A:Introns: 293/1; 342/1
A:Function: promotes resistance to Peronospora parasitica

Query Match 70.0%; Score 35; DB 2; Length 906;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXNDY 10
| : | | | : |
Db 881 EKLVPQGEDY 890

RESULT 6
T48899
disease resistance protein rpp8 [similarity] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 21-Jul-2000
C:Accession: T48899
R:McDowell, J.M.; Dhandaydham, M.; Long, T.A.; Aarts, M.G.; Goff, S.; Holub, E.B.; Plant Cell 10, 1861-1874, 1998
A:Title: Intragenic recombination and diversifying selection contribute to the evolution of the Arabidopsis thaliana rpp8 gene
A:Reference number: Z24999; MUID:99030193; PMID:9811794
A:Accession: T48899
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-908 <MCD>

A:Cross-references: EMBL:AF089711; NID:g3901293; PIDN:AAC78631.1; PID:g3901294
 A:Experimental source: Columbia
 C:Genetics:
 A:Gene: rpp8
 A:Introns: 293/1; 342/1
 C:Function:
 A:Description: susceptible allele of a gene that promotes resistance to Peronospora para

Query Match 70.0%; Score 35; DB 2; Length 908;
 Best Local Similarity 60.0%; Pred. NO. 40;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 I:| | | | |
 Db 883 EKLVPGGDY 892

RESULT 7
 A42452
 V1 protein - tobacco yellow dwarf virus (strain Australia)
 C:Species: tobacco yellow dwarf virus
 C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
 C:Accession: A42452
 R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
 Virolgy 187, 633-642, 1992
 A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow
 A:Reference number: A42452; MUID:92188538; PMID:1546458
 A:Accession: A42452
 A:Molecule type: DNA
 A:Residues: 1-102 <MOR>
 A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 68.0%; Score 34; DB 2; Length 102;
 Best Local Similarity 60.0%; Pred. NO. 6;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXDXS 11
 I:| | | | |
 Db 7 QVVPSSINYS 16

RESULT 8
 S54619
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
 C:Accession: S54619; S66879
 R:de Haan, M.; Maarse, A.C.; Grivell, L.A.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: S54617
 A:Accession: S54619
 A:Molecule type: DNA
 A:Residues: 1-156 <DEH>
 A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123
 R:de Haan, M.; Grivell, L.A.; Maarse, A.C.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66877
 A:Accession: S66879
 A:Molecule type: DNA
 A:Residues: 1-156 <DEW>
 A:Cross-references: EMBL:274920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR01
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0005539
 A:Map position: 15R
 C:Superfamily: hypothetical protein YOR013w

Query Match 68.0%; Score 34; DB 2; Length 156;
 Best Local Similarity 66.7%; Pred. NO. 9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXDX 10

Db 50 EVMPLGMDY 58
 I:| | | | |

RESULT 9
 A96001

conserved hypothetical protein, homolog to osmotically inducible sensory protein S
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: A96001
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.;
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixin
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: A96001
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49673.1; PID:g15141160; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-P
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher,
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Le
 hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Y
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMB21444
 A:Genome: plasmid

Query Match 68.0%; Score 34; DB 2; Length 247;
 Best Local Similarity 60.0%; Pred. NO. 16;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDX 10
 I:| | | | |
 Db 48 EDVEPRGADY 57

RESULT 10
 A96546
 unknown protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96546
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewa
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Mar
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mar
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Ta
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A96546
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <STO>
 A:Cross-references: GB:AE005173; NID:g11094688; PIDN:AAG29624.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F8A12.12
 A:Map position: 1

Query Match 68.0%; Score 34; DB 2; Length 257;
 Best Local Similarity 54.5%; Pred. NO. 17;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDXS 11

```

Db      217 EELLKAGADYS 227
|||||
RESULT 11
F82491
ferrisiderophore reductase VCA0183 [imported] - Vibrio cholerae (strain N16961 serogroup
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82491
Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82491
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <HE>
A:Cross-references: GB:AE004358; GB:AE003853; NID:g9657566; PIDN:AAF96096.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0183
A:Map position: 2
C:Superfamily: flavohemoglobin; cytochrome-b5 reductase homology; globin homology

Query Match      68.0%; Score 34; DB 2; Length 394;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 EVVPXGXDY 10
|||||
Db      194 EVTPGSDY 202
|||||

RESULT 12
H87660
peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87660
R:Nielsen, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87660
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <SFO>
A:Cross-references: GB:AE005673; NID:gl3425020; PIDN:AAK25284.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3322

Query Match      68.0%; Score 34; DB 2; Length 433;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGXDYS 11
|||||
Db      266 EVILPGFDYS 276
|||||

RESULT 13
B49132
fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
C:Accession: B49132; A49132
R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
Development 116, 985-1000, 1992

```

```

A:Title: The fat facets gene is required for Drosophila eye and embryo development
A:Reference number: A49132; MUID:93202020; PMID:1295747
A:Contents: isogenic st
A:Accession: B49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2747 <FIS>
A:Cross-references: GB:104959; NID:gl57411; PIDN:AAF01345.1; PID:g6013474
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:1.
A:Accession: A49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2704, 'VT', 2707, 'ANNV' <FI2>
A:Cross-references: GB:104958; NID:gl57410; PIDN:AAF01346.1; PID:g6013475
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:1.
C:Keywords: alternative splicing

Query Match      68.0%; Score 34; DB 2; Length 2747;
Best Local Similarity 54.5%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGXDYS 11
|||||
Db      1394 EVIVPDGQDFS 1404
|||||

RESULT 14
VKLJST
trans-regulatory splicing protein - simian immunodeficiency virus SIVcpz
N:Alternate names: anti-repression trans-activator; art protein; rev protein; trs
C:Species: simian immunodeficiency virus SIVcpz
A>Note: host Pan troglodytes (chimpanzee)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: S09988
R:Huet, T.; Cheynier, R.; Meyerhans, A.; Roelants, G.; Wain-Hobson, S.
Nature 345, 356-359, 1990
A:Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.
A:Reference number: S09983; MUID:90259077; PMID:2188136
A:Accession: S09988
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-124 <HUE>
A:Cross-references: EMBL:X52154; NID:g58866; PIDN:CAA36405.1; PID:g763085
C:Genetics:
A:Gene: rev; trs; art
A:Introns: 27/1
C:Superfamily: AIDS trans-regulatory splicing protein
C:Keywords: AIDS; immunodeficiency; splicing protein; transcription regulation

Query Match      66.0%; Score 33; DB 1; Length 124;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 EVVPXGXDYS 11
|||||
Db      107 ETVPAGGNYS 116
|||||

RESULT 15
AG1272
thiol peroxidases homolog lmo1583 [imported] - Listeria monocytogenes (strain EGD-
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C:Accession: AG1272
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bli
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsi
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam,
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wei
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

```

A:Accession: AG1272
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-165 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC99661.1; PID:g16411012; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1583
 C:Superfamily: thioredoxin peroxidase

Query Match 66.0%; Score 33; DB 2; Length 165;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10
 |||||
 DB 144 EVVPEGSDH 152

RESULT 16

AH1635
 Thiol peroxidases homolog lin1625 [imported] - *Listeria innocua* (strain Clip11262)
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
 C:Accession: AH1635
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1635
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-165 <GLA>
 A:Cross-references: PIDN:AL592022; PIDN:CAC96856.1; PID:g16414112; GSPDB:GN00178
 A:Experimental source: strain Clip11262
 C:Genetics:
 A:Gene: lin1625
 C:Superfamily: thioredoxin peroxidase

Query Match 66.0%; Score 33; DB 2; Length 165;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGXDY 10
 |||||
 DB 144 EVVPEGSDH 152

RESULT 17

AD0454
 conserved hypothetical protein YPO3732 [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: AD0454
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AD0454
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-196 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93200.1; PID:g15981648; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO3732

Query Match 66.0%; Score 33; DB 2; Length 196;

Best Local Similarity 45.5%; Pred. No. 20;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
 ||:|
 DB 76 EEAIPTSDDYA 86

RESULT 18

S57810
 hypothetical protein precursor (clone TP11) - tomato
 C:Species: *Lycopersicon esculentum* (tomato)
 C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
 C:Accession: S57810
 R:Milligan, S.B.; Gasser, C.S.
 Plant Mol. Biol. 28, 691-711, 1995
 A:Title: Nature and regulation of pistil-expressed genes in tomato.
 A:Reference number: S57808; MUID:95375233; PMID:7647301
 A:Accession: S57810
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-225 <MIL>
 A:Cross-references: EMBL:U20592; NID:9924625; PIDN:AAA80497.1; PID:9924626
 C:Superfamily: plant kunitz-type proteinase inhibitor

Query Match 66.0%; Score 33; Length 225;
 Best Local Similarity 54.5%; Pred. No. 23;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
 :|||
 DB 32 DEVVPGKTYA 42

RESULT 19

S40753
 hypothetical protein C15H7.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
 C:Accession: S40753
 R:Smith, A.
 submitted to the EMBL Data Library, February 1993
 A:Reference number: S40750
 A:Accession: S40753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <SMT>
 A:Cross-references: EMBL:Z22173; NID:g297944; PID:g297948
 C:Genetics:
 A:Introns: 14/3; 59/2; 115/2; 188/2; 238/3

Query Match 66.0%; Score 33; DB 2; Length 327;
 Best Local Similarity 63.6%; Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
 :|||
 DB 175 KEVVPNGDKS 185

RESULT 20

DERTCM
 acyl-CoA dehydrogenase (EC 1.3.99.3) precursor, medium-chain-specific, mitochondrial
 N:Alternate names: acyl dehydrogenase, medium-chain-specific
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 01-Dec-2000
 C:Accession: A28436; S15128
 R:Matsubara, Y.; Kraus, J.P.; Ozasa, H.; Glassberg, R.; Finocchiaro, G.; Ikeda, Y.
 J. Biol. Chem. 262, 10104-10108, 1987
 A:Title: Molecular cloning and nucleotide sequence of cDNA encoding the entire pre
 A:Reference number: A28436; MUID:87280028; PMID:3611054
 A:Accession: A28436
 A:Molecule type: mRNA

A;Residues: 1-421 <RF1>
A;Cross-references: GB:J02791; NID:g202688; PIDN:AAA40670.1; PID:g202689
R;Inagaki, T.; Ohishi, N.; Tsukagoshi, N.; Uda, S.; Ghisla, S.; Yagi, K.
Biochim. Biophys. Acta 1077, 285-290, 1991
A;Title: Structurally different rat liver medium-chain acyl CoA dehydrogenases directed
A;Reference number: S15128; MUID:91230137; PMID:2029527
A;Accession: S15128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 11-81 <BIO>
C;Superfamily: acyl-CoA dehydrogenase
C;Keywords: fatty acid beta-oxidation; fatty acid metabolism; flavoprotein; mitochondrion
F;1-25/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F;26-421/Product: acyl-CoA dehydrogenase, medium-chain-specific #status predicted <MAT>

Query Match 66.0%; Score 33; DB 1; Length 421;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPKXGXDY 10
||:| | |
DB 58 EEIIPVADPY 67

RESULT 21
H72784
Probable alkaline proteinase APE0263 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: H72784
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A7450; MUID:99310339; PMID:10382966
A;Accession: H72784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-440 <KAW>
A;Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79178.1; PID:g5103657
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0263
C;Superfamily: subtilisin; subtilisin homology

Query Match 66.0%; Score 33; DB 2; Length 440;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPKXGXDY 10
||:| | |
DB 120 EVLPWGVVDY 128

RESULT 22
AF3286
ATP-dependent DNA helicase BMEI0275 [Imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AF3286
R;DeiVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujeer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AF3286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1028 <KUR>
A;Cross-references: GB:AF008917; PIDN:AAL51457.1; PID:g17982167; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0275

A;Map position: I

Query Match 66.0%; Score 33; DB 2; Length 1028;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPKXGXDYS 11
||:| | |
DB 76 EKIVPPGARYS 86

RESULT 23

D82246
Probable chitinase VC1073 [Imported] - Vibrio cholerae (strain N16961 serogroup O1
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: D82246
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson,
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sell
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82246
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1088 <HEI>
A;Cross-references: GB:AE004188; GB:AE003852; NID:g9655530; PIDN:AAF94232.1; GSPDB:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1073
A;Map position: 1

Query Match 66.0%; Score 33; DB 2; Length 1088;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVVPKXGXDYS 11
||:| | |
DB 966 ETIPAGGEYS 975

RESULT 24

T20173
Hypothetical protein C53A5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T20173; T23857
R;Mortimore, B.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19232
A;Accession: T20173
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1150 <WIL>
A;Cross-references: EMBL:Z81486; PIDN:CAB03994.1; GSPDB:GN00023; CESP:C53A5.2
A;Experimental source: clone C53A5
R;Matthews, L.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19808
A;Accession: T23857
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1150 <WT2>
A;Cross-references: EMBL:Z78015; PIDN:CAB01437.1; GSPDB:GN00023; CESP:C53A5.2
A;Experimental source: clone R02D5
C;Genetics:
A;Gene: CESP:C53A5.2
A;Map position: 5
A;Introns: 33/3; 63/3; 132/3; 169/3; 221/3; 299/3; 379/2; 423/2; 438/2; 471/1; 513/

Query Match 66.0%; Score 33; DB 2; Length 1150;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VWPXGXDYS 11

I: | | | |

Db 562 VLPVGIDYS 570

RESULT 25

S00210

plastocyanin b - Lombardy poplar

C:Species: Populus nigra var. italica (Lombardy poplar)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 20-Apr-2000

C:Accession: S00210

R:Dimitrov, M.I.; Egorov, C.A.; Donchev, A.A.; Atanasov, B.P.

FEBS Lett. 226, 17-22, 1987

A:Title: Complete amino acid sequence of poplar plastocyanin b.

A:Reference number: S00210

A:Accession: S00210

A:Molecule type: protein

A:Residues: 1-99 <blm>

C:Superfamily: plastocyanin

C:Keywords: Chloroplast; copper; electron transfer; metalloprotein

F:37,84,87,92/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match

64.0%; Score 32; DB 2; Length 99;

Best Local Similarity 54.5%; Pred. No. 16;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVWPXGXDYS 11

I: | | | |

Db 43 EDAPSGVDVS 53

Search completed: June 10, 2003, 13:49:18

Job time : 11.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-46
Perfect score: 50
Sequence: 1 EEVVPXGXDYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	70.0	1058	1	CARB_FUSNN
2	34	68.0	102	1	YLIK_TYDVA
3	34	68.0	394	1	HMPA_VIBCH
4	34	68.0	2747	1	FAF_DROME
5	33	66.0	124	1	REV_SIVCZ
6	33	66.0	327	1	YK14_CAEEL
7	33	66.0	421	1	ACDM_RAT
8	32	64.0	154	1	PLAS_ORYSA
9	32	64.0	155	1	PLAS_HORVU
10	32	64.0	168	1	PLAT_POPNI
11	32	64.0	196	1	XJAG_SALTY
12	32	64.0	427	1	TOUB_HAETN
13	32	64.0	1499	1	ALOC_HUMAN
14	32	64.0	3174	1	CHAC_HUMAN
15	31	62.0	233	1	HIS9_THEMA
16	31	62.0	319	1	YHAI_CRYPA
17	31	62.0	421	1	ACDM_HUMAN
18	31	62.0	421	1	ACDM_PIG
19	31	62.0	562	1	TR2M_ERWHE
20	30	60.0	97	1	PLAS_DAUCA
21	30	60.0	121	1	TKNK_HUMAN
22	30	60.0	175	1	HES3_RAT
23	30	60.0	203	1	CTC_BACSU
24	30	60.0	231	1	ARAD_ECOLI
25	30	60.0	231	1	ARAD_SALTY
26	30	60.0	232	1	SCOA_HELPJ
27	30	60.0	232	1	SCOA_HELPY
28	30	60.0	259	1	OVUH_LYMST
29	30	60.0	421	1	AMP2_YEAST
30	30	60.0	421	1	ECB2_HALEL
31	30	60.0	423	1	ECB1_HALEL
32	30	60.0	430	1	FOLC_BACSU
33	30	60.0	457	1	Z185_HUMAN

34	30	60.0	469	1	LET1_KLULA
35	30	60.0	478	1	AMP2_HUMAN
36	30	60.0	478	1	AMP2_MOUSE
37	30	60.0	478	1	AMP2_RAT
38	30	60.0	478	1	GSR2_HUMAN
39	30	60.0	530	1	APPF_XANCH
40	30	60.0	556	1	APLY_APLKU
41	30	60.0	583	1	HEMO_BRARE
42	30	60.0	890	1	BCN5_CLOPE
43	30	60.0	1049	1	SPS_ORYSA
44	30	60.0	1068	1	SPS_MAIZE
45	30	60.0	1081	1	SPS2_CRAPL
46	30	60.0	1176	1	NIR_NEUCR
47	30	60.0	1401	1	RPOC_VIBCH
48	30	60.0	1501	1	SNO2_YEAST
49	30	60.0	1729	1	TABP_HUMAN
50	30	60.0	2717	1	ZEP1_HUMAN
51	30	60.0	3099	1	POLG_PEMVM
52	29.5	59.0	472	1	ET2A_XENLA
53	29	58.0	98	1	PLAS_ENTPR
54	29	58.0	98	1	PLAS_ULVAR
55	29	58.0	98	1	PLAS_ULVPE
56	29	58.0	99	1	PLAS_RUMOB
57	29	58.0	99	1	PLAS_TOBAC
58	29	58.0	113	1	Y011_BPT4
59	29	58.0	175	1	HES3_MOUSE
60	29	58.0	230	1	YCAP_ECOLI
61	29	58.0	242	1	PSB4_XENLA
62	29	58.0	258	1	YC10_METJA
63	29	58.0	261	1	KLKD_MOUSE
64	29	58.0	338	1	MTBA_METBA
65	29	58.0	354	1	VGLI_VZVD
66	29	58.0	388	1	FEM3_CAEEL
67	29	58.0	389	1	SAT_AERPE
68	29	58.0	412	1	DOK2_HUMAN
69	29	58.0	421	1	ACDM_MOUSE
70	29	58.0	424	1	MS11_ARATH
71	29	58.0	424	1	MS11_LYCES
72	29	58.0	435	1	HEM1_DESYM
73	29	58.0	471	1	TMLH_NEUCR
74	29	58.0	472	1	ET2B_XENLA
75	29	58.0	477	1	CAP2_RAT

ALIGNMENTS

RESULT 1

ID	CARB_FUSNN	STANDARD	PRT	1058 AA
AC	Q8RG86			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).			
GN	CARB OR FN0422.			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21886394; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyripides N., Overbeek R.			
RA	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586."			
RT	J. Bacteriol. 184:2005-2018(2002).			
CC	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			

phosphate + L-glutamate + carbamoyl phosphate.
 -!- COPACTOR: Binds three manganese ions (By similarity).
 -!- PATHWAY: Arginine biosynthesis.
 -!- SUBUNIT: Pyrimidine biosynthesis; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).
 -!- SIMILARITY: BELONGS TO THE CARB FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE010554; AAL94625.1; ALT_INIT.
 InterPro; IPR005483; CPase_L.
 InterPro; IPR005479; CPase_L_D2.
 InterPro; IPR005480; CPase_L_D3.
 InterPro; IPR005481; CPase_L_N.
 InterPro; IPR004362; MGS_like.
 Pfam; PF00289; CPase_L_chain; 2.
 Pfam; PF02786; CPase_L_D2; 2.
 Pfam; PF02787; CPase_L_D3; 1.
 Pfam; PF02142; MGS; 1.
 PRINTS; PR00098; CPASE.
 PROSITE; PS00866; CPASE_1; 2.
 PROSITE; PS00867; CPASE_2; 2.
 Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 ATP-binding; Manganese; Complete proteome.
 DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 DOMAIN 930 1058 ALLOSTERIC DOMAIN.
 REPEAT 1 546
 REPEAT 547 1058
 NP_BIND 153 210 ATP (POTENTIAL).
 NP_BIND 302 352 ATP (POTENTIAL).
 METAL 284 298 MANGANESE 1 (BY SIMILARITY).
 METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
 METAL 300 300 MANGANESE 2 (BY SIMILARITY).
 METAL 820 820 MANGANESE 3 (BY SIMILARITY).
 METAL 832 832 MANGANESE 3 (BY SIMILARITY).
 SEQUENCE 1058 AA; 117451 MW; ED7037AF7C1E339f CRC64;

Query Match 70.0%; Score 35; DB 1; Length 1058;
 Best Local Similarity 60.0%; Pred. No. 17;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGDYDS 11
 : : : : :
 Db 190 EIVPGLNYS 199

RESULT 2
 Y11K_TYDVA
 ID Y11K_TYDVA STANDARD; PRT; 102 AA.
 AC P31619;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE Hypothetical 11.2 kDa protein.
 OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
 GN Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
 OC NCBI_TaxID=31599;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92188538; PubMed=1546458;
 RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

"The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants.";
 Virology 187:633-642(1992).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M81103; AAA47947.1; -.
 PIR; A42452;
 InterPro; IPR002621; Gemini_mov.
 DR Pfam; PF01708; Gemini_mov; 1.
 KW Hypothetical protein
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 102;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGDYDS 11
 : : : : :
 Db 7 QVVPNGINYS 16

RESULT 3
 HMPA_VIBCH
 ID HMPA_VIBCH STANDARD; PRT; 394 AA.
 AC Q9KMY3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
 GN HMP OR VCA0183.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=El Tor N16961 / Serotype O1;
 RC MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Peterson J.D., Umayam L.A., Dodson R.J., Haft D.H., Hickey E.K., Tettelin H., Richardson D., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Dragoi I., Sellers P., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Salzman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae".
 RL Nature 406:477-483(2000).
 CC -!- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS; ONE IS A HEME-CONTAINING OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
 CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN FLAVOHEMOPROTEINS SUBFAMILY.
 CC -!- SIMILARITY: TO A NUMBER OF FAD/NAD(P) FLAVOPROTEIN OXIDOREDUCTASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE004358; AAF96096.1; -.
 HSP; P39662; 1CQX.

TIGR; VCA0183; -
 DR InterPro; IPR001834; Cyt B5 reductase.
 DR InterPro; IPR001709; FPN_cyt_reductase.
 DR InterPro; IPR000971; Globin.
 DR InterPro; IPR001433; Oxid_FAD/NAD(P).
 DR InterPro; IPR001221; Phe_hydroxylase.
 DR Pfam; PF00042; globin; 1.
 DR Pfam; PF00175; NAD_binding; 1.
 DR Pfam; PF00970; FAD_binding_6; 1.
 DR PRINTS; PR00371; FPNCR.
 DR PRINTS; PR00410; PHEHYDROLASE.
 DR PROSITE; PS01033; GLOBIN; 1.
 KW Oxidoreductase; NADP; Heme; Flavoprotein; FAD; Iron transport;
 FT Oxygen transport; Transport; Complete proteome.
 FT DOMAIN 1 136
 FT METAL 53 53
 FT METAL 85 85
 FT METAL 85 85
 FT NP_BIND 268 273
 FT SEQUENCE 394 AA; 44191 MW; DDA3490FAE28823A CRC64;
 Query Match 68.0%; Score 34; DB 1; Length 394;
 Best Local Similarity 66.7%; Pred. No. 9.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 EVVPXGXDY 10
 Db 194 EVTPGSDY 202
 |||||
 RESULT 4
 FAF_DROME STANDARD; PRT; 2747 AA.
 AC P55824;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
 DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF)
 DE (Deubiquitinating enzyme FAF) (Fat facets protein).
 GN FAF.
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93202020; PubMed=1295747;
 RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;
 RT "The fat facets gene is required for Drosophila eye and embryo development."
 RL Development 116:985-1000(1992).
 CC -!- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A
 CC ROLE IN COMPOUND EYE ASSEMBLY AND OGENESIS RESPECTIVELY. IN THE
 CC LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS
 CC PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY
 CC CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR
 CC NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND
 CC COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR
 CC FUNCTION.
 CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; L04959; AAF01345.1; -
 CC EMBL; L04958; AAF01346.1; -
 CC MEROPS; C19.007; -
 DR FlyBase; FBgn0005632; faf.
 DR InterPro; IPR001394; UCH-2.
 DR Pfam; PF00442; UCH-1; 1.
 DR Pfam; PF00443; UCH-2; 1.
 DR PROSITE; PS00972; UCH_2_1; 1.
 DR PROSITE; PS00973; UCH_2_2; 1.
 DR PROSITE; PS0235; UCH_2_3; 1.
 KW Ubl conjugation pathway; Hydrolase; Thiol protease;
 KW Developmental protein; Vision; Alternative splicing.
 FT ACT_SITE 1677 1677
 FT ACT_SITE 1978 1978
 FT ACT_SITE 1986 1986
 FT VARSPLIC 2705 2747
 FT VARIANT 2725 2725
 FT SEQUENCE 2747 AA; 307954 MW; ID97659F7A7B2ADE CRC64;
 Query Match 68.0%; Score 34; DB 1; Length 2747;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGXDSY 11
 Db 1394 EEVVPDGDGDFS 1404
 |||||
 RESULT 5
 REV_SIVCZ STANDARD; PRT; 124 AA.
 AC P17280;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE REV protein (Anti-repression transactivator protein) (ART/TRS).
 DE REV.
 GN REV.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90259077; PubMed=2188136;
 RA Huet T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
 RT "Genetic organization of a chimpanzee lentivirus related to HIV-1."
 RL Nature 345:356-359(1990).
 CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
 CC -!- NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
 CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
 CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
 CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X52154; CAA36405.1; -
 CC PIR; S09888; VKLJSI.
 CC HIV; X52154; REVSCPZ.
 DR InterPro; IPR000625; REV_protein.
 DR Pfam; PF00424; REV; 1.
 KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
 SQ SEQUENCE 124 AA; 13701 MW; F5877D1BDF65A7B2 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 124;
 Best Local Similarity 60.0%; Pred. No. 4.6;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EWPVPGXDYS 11
 Db 107 ETVPAGGNYS 116

RESULT 6

YK14_CAEEL STANDARD; PRT; 327 AA.
 AC P34338;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C15H7.4 in chromosome III.
 GN C15H7.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Fraser A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaughn M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z22173; CAA80126.1; -;
 DR PIR; S40753; S40753.
 DR WormPep; C15H7.4; CE00082.
 KW Hypothetical protein.
 SQ SEQUENCE 327 AA; 35566 MW; 716BC2BDD2E9607E CRC64;

Query Match 66.0%; Score 33; DB 1; Length 327;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDYS 11
 Db 175 KEVVPNGGDKS 185

RESULT 7

ACDM_RAT STANDARD; PRT; 421 AA.
 AC P08503;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
 DE (EC 1.3.99.3) (NCAD).

GN ACADM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=87280028; PubMed=3611054;
 RA Matsubara Y., Kraus J.P., Ozasa H., Glassberg R., Finocchiaro G.,
 RA Ikeda Y., Mole J., Rosenberg L.E., Tanaka K.;
 RT "Molecular cloning and nucleotide sequence of cDNA encoding the
 RT entire precursor of rat liver medium chain acyl coenzyme A
 RT dehydrogenase.";
 RL J. Biol. Chem. 262:10104-10108(1987).
 CC -1- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO
 CC 16.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
 CC step.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
 CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
 CC TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02791; AAA40670.1; -;
 DR PIR; A28436; DERTCM.
 DR HSSP; P11310; LEGD.
 DR InterPro; IPR001352; Acyl-CoA_dh.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
 DR PROSITE; PS00072; Acyl-CoA_DH_1; 1.
 DR PROSITE; PS00073; Acyl-CoA_DH_2; 1.
 KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
 KW Mitochondrion; Transient peptide.
 FT TRANSIT 1 25 MITOCHONDRION.
 FT CHAIN 26 421 ACYL-COA DEHYDROGENASE, MEDIUM-CHAIN
 FT SPECIFIC.
 FT ACT_SITE 193 193 FORMS A HYDROGEN-BOND WITH THE FLAVIN
 FT ACT_SITE 401 401 N(5) OF THE FAD COFACTOR (BY SIMILARITY).
 FT SEQUENCE 421 AA; 46555 MW; 2CF076F8C919BDE8 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 421;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 Db 58 EEIIPVADY 67

RESULT 8
 ID PLAS_ORISA STANDARD; PRT; 154 AA.
 AC P20423; Q9SBB8;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin, chloroplast precursor.

GN PETE.
 OS *Oryza sativa* (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; *Oryza*.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Ilpoom; TISSUE=Leaf;
 RA Lee J.-S.;
 RT "Molecular cloning and characterization of plastocyanin precursor in
 rice.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 58-154.
 RC STRAIN=cv. Japonica;
 RX MEDLINE=89386623; PubMed=2780537;
 RA Yano H., Kano M., Tsugita A., Aso K., Nozu Y.;
 RT "The amino acid sequence of plastocyanin from rice (*Oryza sativa*,
 subspecies japonica).";
 RL Protein Seq. Data Anal. 2:385-389(1989).
 CC -!- FUNCTION: Participates in electron transfer between P700 and the
 cytochrome b6-f complex in photosystem I.
 CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF093636; AAC78108.1; -.
 DR PIR; S06105; S06105.
 DR PIR; JT0352; JT0352.
 DR HSP; P00289; 2PCF.
 DR InterPro; IPR000923; BlueCu_1.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PD00156; COPPERBLUE.
 DR ProDom; PD001235; Copper_blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 DR Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide.
 FT CHAIN 1 57 CHLOROPLAST.
 FT DOMAIN 58 154 PLASTOCYANIN.
 FT METAL 94 94 COPPER (BY SIMILARITY).
 FT METAL 139 139 COPPER (BY SIMILARITY).
 FT METAL 142 142 COPPER (BY SIMILARITY).
 FT METAL 147 147 COPPER (BY SIMILARITY).
 SQ SEQUENCE 154 AA; 15577 MW; E45725D2B5F400D CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 154;
 Best Local Similarity 54.5%; Pred. No. 9.6;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db 100 EDAPVSGVDVS 110
 RESULT 9
 ID PLAS_HORVU STANDARD; PRT; 155 AA.
 AC P08248;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin, chloroplast precursor.

GN PETE.
 OS *Hordeum vulgare* (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Triticeae; *Hordeum*.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Boml;
 RA Nielsen O.S., Gausling K.;
 RT "The precursor of barley plastocyanin: sequence of cDNA clones and
 gene expression in different tissues.";
 RL FEBS Lett. 225:159-162(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. NK 1558;
 RX MEDLINE=94039081; PubMed=8223592;
 RA Nielsen P., Gausling K.;
 RT "In vitro binding of nuclear proteins to the barley plastocyanin gene
 promoter region.";
 RL Eur. J. Biochem. 217:97-104(1993).
 CC -!- FUNCTION: Participates in electron transfer between P700 and the
 cytochrome b6-f complex in photosystem I.
 CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y00704; CAA68696.1; -.
 DR EMBL; Z28347; CAA82201.1; -.
 DR PIR; S00206; S00206.
 DR HSP; P00289; 2PCF.
 DR InterPro; IPR000923; BlueCu_1.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PD00156; COPPERBLUE.
 DR ProDom; PD001235; Copper_blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 DR Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide.
 FT CHAIN 1 58 CHLOROPLAST.
 FT DOMAIN 59 155 PLASTOCYANIN.
 FT METAL 95 95 COPPER (BY SIMILARITY).
 FT METAL 140 140 COPPER (BY SIMILARITY).
 FT METAL 143 143 COPPER (BY SIMILARITY).
 FT METAL 148 148 COPPER (BY SIMILARITY).
 FT VARIANT 120 120 T -> N (IN CV. NK 1558).
 SQ SEQUENCE 155 AA; 15709 MW; DAA7EAB5F6F4F91 CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 155;
 Best Local Similarity 54.5%; Pred. No. 9.6;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDYS 11
 Db 101 EDAPVSGVDVS 111
 RESULT 10
 ID PLAT_POPNI STANDARD; PRT; 168 AA.
 AC P11970;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Plastocyanin B, chloroplast precursor.
 GN PEST.
 OS Populus nigra (Lombardy poplar).
 ON Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Italica; TISSUE=Leaf;
 RA Reichert J., Jenzelewski V., Haehnel W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 70-168.
 RC STRAIN=cv. Italica;
 RA Dmitrov M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;
 RT FEBS Lett. 226:17-22(1987).
 RL
 CC -1- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF
 CC POPULAR PLASTOCYANINS A AND B.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z50186; CAA90565.1; -
 CC PIR; S00210; S00210.
 CC HSP; P00299; 1PLC.
 CC InterPro; IPR000923; BlueCu1.
 CC InterPro; IPR001235; Copper_blue.
 CC Pfam; PF00127; copper-bind; 1.
 CC PRINTS; PR00156; COPPERBLUE.
 CC ProDom; PD001235; Copper blue; 1.
 CC PROSITE; PS00196; COPPER_BLUE; 1.
 CC Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide; Multigene family.
 FT TRANSIT 1 69
 FT CHAIN 70 168
 FT DOMAIN 70 168
 FT PLASTOCYANIN B.
 FT PLASTOCYANIN-LIKE.
 FT METAL 106 106
 FT METAL 153 153
 FT METAL 156 156
 FT METAL 161 161
 FT COPPER.
 FT COPPER.
 SQ SEQUENCE 168 AA; 16981 MW; F20DA6EA2038AEEA CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 168;
 Best Local Similarity 54.5%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGNDYS 11
 Db 112 EDAPSGVDVS 122
 ID YVAG_SALTY STANDARD; PRT; 196 AA.
 AC Q9L9I2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yJag.
 GN YJAG OR STM4169 OR STM1.23 OR STV3716.
 OS Salmonella typhimurium, and

OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE HI0431.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF170176; AAF33519.1; -
 CC EMBL; AE008894; AAL22997.1; -
 CC EMBL; AL627279; CAD09475.1; -
 CC StyGene; SG????; yJag.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 196 AA; 22704 MW; E40049CE5EE64150 CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 196;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 Db 76 EEAI PAADY 85
 ID TOLB_HAEIN STANDARD; PRT; 427 AA.
 AC P44677; P94811;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE TOLB protein precursor.
 GN TOLB OR HI0382.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Rd.";
 RL Science 269:496-512(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1479;
 RA MEDLINE=97080550; PubMed=8921895;
 RA Sen K., Sikkema D.J., Murphy T.F.;
 RT "Isolation and characterization of the Haemophilus influenzae tolQ,
 RT tolR, tolA and tolB genes.";
 RL Gene 178:75-81(1996).
 CC -!- FUNCTION: INVOLVED IN THE TONB-INDEPENDENT UPTAKE OF PROTEINS
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TOLB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U32722; AAC22040.1; -;
 CC DR EMBL; U32470; AAC44597.1; -;
 CC DR HSSP; P19935; 1CRZ.
 CC TIGR: HI0382; -;
 KW Transport; Protein transport; Periplasmic; Signal; Complete proteome.
 FT SIGNAL 1 23
 FT CHAIN 24 427
 FT VARIANT 6 6 R -> H (IN STRAIN 1479).
 FT VARIANT 14 14 V -> I (IN STRAIN 1479).
 FT VARIANT 17 19 VGS -> ITH (IN STRAIN 1479).
 FT VARIANT 21 21 A -> V (IN STRAIN 1479).
 FT VARIANT 79 79 R -> H (IN STRAIN 1479).
 FT VARIANT 129 129 T -> A (IN STRAIN 1479).
 FT VARIANT 160 160 A -> G (IN STRAIN 1479).
 FT VARIANT 237 237 A -> T (IN STRAIN 1479).
 FT VARIANT 322 322 S -> N (IN STRAIN 1479).
 FT VARIANT 326 326 A -> V (IN STRAIN 1479).
 FT VARIANT 328 328 A -> S (IN STRAIN 1479).
 SQ SEQUENCE 427 AA; 44967 MW; 0882201AE5254B9 CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 427;
 Best Local Similarity 60.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 EVVPGXDYS 11
 Db 103 QVVPNGGYS 112
 RESULT 13
 ID A10C.HUMAN STANDARD; PRT; 1499 AA.
 AC O60312: Q96914;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)
 DE (Aminophospholipid translocase VC).
 GN ATP10C OR ATPVC OR KIAA0566.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21225279; PubMed=11326269;
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
 RA Oshimura M.;
 RT "A novel maternally expressed gene, ATP10C, encodes a putative
 RT aminophospholipid translocase associated with Angelman syndrome.";
 RL Nat. Genet. 28:19-20(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21313119; PubMed=11353404;
 RA Herzig L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;
 RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
 RT adjacent to UBE3A and exhibits similar imprinted expression.";
 RL Am. J. Hum. Genet. 68:1501-1505(2001).
 RN [3]
 RP SEQUENCE OF 337-1499 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Wide expression, with highest levels in
 CC kidney, followed by lung, brain, prostate, testis, ovary, and
 CC small intestine.
 CC -!- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome
 CC (AS), also known as 'happy puppet syndrome'.
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
 CC ATPASES). SUBFAMILY IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AB051358; BAB47392.1; -;
 CC DR EMBL; AY029504; AAK33100.1; -;
 CC DR EMBL; AY029487; AAK33100.1; JOINED.
 CC DR EMBL; AY029488; AAK33100.1; JOINED.
 CC DR EMBL; AY029489; AAK33100.1; JOINED.
 CC DR EMBL; AY029490; AAK33100.1; JOINED.
 CC DR EMBL; AY029491; AAK33100.1; JOINED.
 CC DR EMBL; AY029492; AAK33100.1; JOINED.
 CC DR EMBL; AY029493; AAK33100.1; JOINED.
 CC DR EMBL; AY029494; AAK33100.1; JOINED.
 CC DR EMBL; AY029495; AAK33100.1; JOINED.
 CC DR EMBL; AY029496; AAK33100.1; JOINED.
 CC DR EMBL; AY029497; AAK33100.1; JOINED.
 CC DR EMBL; AY029498; AAK33100.1; JOINED.
 CC DR EMBL; AY029499; AAK33100.1; JOINED.
 CC DR EMBL; AY029500; AAK33100.1; JOINED.
 CC DR EMBL; AY029501; AAK33100.1; JOINED.
 CC DR EMBL; AY029502; AAK33100.1; JOINED.
 CC DR EMBL; AY029503; AAK33100.1; JOINED.
 CC DR EMBL; AB011138; BAA25492.1; -;
 CC DR Gene; HGNC:13547; ATP10C.
 CC MIM; 605855; -;
 CC MIM; 105830; -;
 CC InterPro; IPR001757; ATPase_E1-E2.
 CC InterPro; IPR001454; Hlgase/hydrolase.
 CC Pfam; PF00702; Hydrolase; 1.


```
FT CONFLICT 2127 2129 GIE -> VSA (IN REF. 4).
FT CONFLICT 2281 2281 G -> E (IN REF. 5).
FT CONFLICT 2413 2413 T -> R (IN REF. 5).
FT CONFLICT 2567 2567 K -> E (IN REF. 5).
SQ SEQUENCE 3174 AA; 360299 MW; FD2AB895B4A7149E CRC64;

Query Match
Best Local Similarity 64.0%; Score 32; DB 1; Length 3174;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
Db 242 ENIVPEGYDF 251

RESULT 15
HIS9_THEME
ID HIS9_THEME STANDARD; PRT; 233 AA.
AC Q9WZRL;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable histidinol-phosphatase (EC 3.1.3.15) (HolPase).
GN HISK OR TM0804.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback F.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Cotton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: L-histidinol-phosphate + H(2)O = L-histidinol
CC + phosphate.
CC -!- PATHWAY: Histidine biosynthesis; eighth step.
CC -!- SIMILARITY: BELONGS TO THE PHP FAMILY OF HYDROLASE. HISK FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
RESULT 16
YHAI_CRYPA
ID YHAI_CRYPA STANDARD; PRT; 319 AA.
AC P10941;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein 1 in hypovirulence-associated DS-RNA genetic
DE element [Contains: P29 proteinase].
DE Cryptonectria parasitica (Chesnut blight fungus) (Endothia
DE parasitica).
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Diaporthales; Valsaceae; Cryptonectria.
OX NCBI_TaxID=5116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=EP713;
RX MEDLINE=89251594; PubMed=2721496;
RA Rae B.P., Hillman B.I., Tartaglia J., Nuss D.L.;
RT "Characterization of double-stranded RNA genetic elements associated
RT with biological control of chestnut blight: organization of terminal
RT domains and identification of gene products.";
RL EMBO J. 8:657-663(1989).
CC -!- MISCELLANEOUS: DOUBLE-STRANDED RNA GENETIC ELEMENTS ARE ASSOCIATED
CC WITH BIOLOGICAL CONTROL OF THE FUNGAL DISEASE CHESTNUT BLIGHT.
CC THIS DS-RNA ARE ASSOCIATED WITH HYPOVIRULENCE. THEY ARE LOCALIZED
CC IN THE CYTOPLASM.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C7.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; X14524; CAA32666.1; -
DR PIR; S03833; S03833.
DR MEROPS; C07.001; -
DR InterPro; IPR002704; Peptidase_C7.
DR Pfam; PF01830; Peptidase_C7; 1.
DR ProDom; PD040949; Peptidase_C7; 1.
DR Hypothetical protein; Hydrolase; Thiol protease.
SQ SEQUENCE 319 AA; 35443 MW; ALF5F775F9AEC7A CRC64;
```

```
Query Match 62.0%; Score 31; DB 1; Length 319;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 EEVVPXG 7
Db 31 EEVVPAG 37
```

```
RESULT 17
ACDM_HUMAN
ID ACDM_HUMAN STANDARD; PRT; 421 AA.
AC P11310; Q9NFF1;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
DE (EC 1.3.99.3) (MCAD).
GN ACDM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=87231952; PubMed=3035565;
RA Kelly D.P., Kim J.-J.P., Billadello J.J., Hainline B.E., Chu T.W.,
```

- RA Strauss A.W.;
RT "Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and
its expression in enzyme-deficient human tissue.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4068-4072(1987).
RN [2]
RC SEQUENCE FROM N.A.
RP TISSUE=Colon;
RA Sun F., Wang Y., Block G.D.;
RT "Medium-chain acyl-CoA dehydrogenase.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE OF 314-342 FROM N.A., AND VARIANT MCAD GLU-329.
RA Matsubara Y., Narisawa K., Miyabayashi S., Tada K., Coates P.M.,
Bachmann C., Elsas L.J. II, Pollitt R.J., Rhead W.J., Roe C.R.;
RT "Identification of a common mutation in patients with medium-chain
acyl-CoA dehydrogenase deficiency.";
RL Biochem. Biophys. Res. Commun. 171:498-505(1990).
RN [5]
RC X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RA Lee H.J., Wang M., Paschke R., Nandy A., Ghisla S., Kim J.J.;
RT "Crystal structures of the wild type and the Glu376Gly/Thr255Glu
mutant of human medium-chain acyl-CoA dehydrogenase: influence of the
location of the catalytic base on substrate specificity.";
RL Biochemistry 35:12412-12420(1996).
RN [6]
RC REVIEW ON VARIANTS MCAD.
RA Tanaka K., Yokota I., Coates P.M., Strauss A.W., Kelly D.P.,
Zhang Z.F., Gregersen N., Andresen B.S., Matsubara Y., Curtis D.,
Chen Y.-T.;
RT "Mutations in the medium chain acyl-CoA dehydrogenase (MCAD) gene.";
RL Hum. Mutat. 1:271-279(1992).
RN [7]
RC VARIANT MCAD GLU-329.
RA MEDLINE-90368980; PubMed-23948825;
Yokota I., Indo Y., Coates P.M., Tanaka K.;
RT "Molecular basis of medium chain acyl-coenzyme A dehydrogenase
deficiency. An A to G transition at position 985 that causes a
lysine-304 to glutamate substitution in the mature protein is the
single prevalent mutation.";
RL J. Clin. Invest. 86:1000-1003(1990).
RN [8]
RC VARIANT MCAD GLU-329.
RA MEDLINE-91067682; PubMed-2251268;
Kelly D.P., Whelan A.J., Ogden M.L., Alpers R., Zhang Z.F., Bellus G.,
Gregersen N., Dorland L., Strauss A.W.;
RT "Molecular characterization of inherited medium-chain acyl-CoA
dehydrogenase deficiency.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9236-9240(1990).
RN [9]
RC VARIANTS MCAD ILE-149; ARG-244; ARG-267 AND THR-375.
RA MEDLINE-92081773; PubMed-1664086;
Yokota I., Coates P.M., Hale D.E., Rinaldo P., Tanaka K.;
RT "Molecular survey of a prevalent mutation, 985A-to-G transition, and
identification of five infrequent mutations in the medium-chain acyl-
CoA dehydrogenase (MCAD) gene in 55 patients with MCAD deficiency.";
RL Am. J. Hum. Genet. 49:1280-1291(1991).
RN [10]
RC VARIANT MCAD GLU-329.
RA MEDLINE-91224627; PubMed-1902818;
Gregersen N., Andresen B.S., Bross P., Winter V., Ruediger N.,
Engst S., Christensen E., Kelly D., Strauss A.W., Koelvræ S.,
Bolund L., Ghisla S.;
RT "Molecular characterization of medium-chain acyl-CoA dehydrogenase
(MCAD) deficiency: identification of a lys329 to glu mutation in the
MCAD gene, and expression of inactive mutant enzyme protein in E.
- RT coli.";
- RL Hum. Genet. 86:545-551(1991).
RN [11]
RC VARIANT MCAD GLU-329 FREQUENCY.
RA MEDLINE-91109464; PubMed-1671131;
Blakemore A.I., Singleton H., Pollitt R.J., Engel P.C., Kolvræ S.,
Gregersen N., Curtis D.;
RT "Frequency of the G985 MCAD mutation in the general population.";
RL Lancet 337:298-299(1991).
RN [12]
RC VARIANTS MCAD THR-326 AND ARG-336.
RA MEDLINE-94256509; PubMed-8198141;
Andresen B.S., Jensen T.G., Bross P., Knudsen I., Winter V.,
Koelvræ S., Bolund L., Ding J.-H., Chen Y.-T., van Hove J.L.K.,
Curtis D., Yokota I., Tanaka K., Kim J.-J.P., Gregersen N.;
RT "Disease-causing mutations in exon 11 of the medium-chain acyl-CoA
dehydrogenase gene.";
RL Am. J. Hum. Genet. 54:975-988(1994).
RN [13]
RC VARIANT MCAD I15-GLY-CYS-116 DEL.
RA MEDLINE-95327425; PubMed-7603790;
Ziaadeh R., Hoffman E.P., Finegold D.N., Hoop R.C., Brackett J.C.,
Strauss A.W., Naylor E.W.;
RT "Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania:
neonatal screening shows high incidence and unexpected mutation
frequencies.";
RL Pediatr. Res. 37:675-678(1995).
RN [14]
RC VARIANT MCAD ARG-195.
RA MEDLINE-95015014; PubMed-7929823;
Brackett J.C., Sims H.F., Steiner R.D., Nunge M., Zimmerman E.M.,
Demartinville B., Rinaldo P., Slaughter R., Strauss A.W.;
RT "A novel mutation in medium chain acyl-CoA dehydrogenase causes sudden
neonatal death.";
RL J. Clin. Invest. 94:1477-1483(1994).
RN [15]
RC VARIANT MCAD ALA-193.
RA MEDLINE-97301766; PubMed-9158144;
Andresen B.S., Bross P., Udvari S., Kirk J., Gray G., Kmoch S.,
Chamoles N., Knudsen I., Winter V., Wilcken B., Yokota I., Hart K.,
Packman S., Harpey J.P., Saudubray J.M., Hale D.E., Bolund L.,
Koelvræ S., Gregersen N.;
RT "The molecular basis of medium-chain acyl-CoA dehydrogenase (MCAD)
deficiency in compound heterozygous patients: is there correlation
between genotype and phenotype?";
RL Hum. Mol. Genet. 6:695-707(1997).
RN [16]
RC CHARACTERIZATION OF VARIANT MCAD ALA-193.
RA MEDLINE-99102080; PubMed-9882619;
Kuchler B., Abdel-Ghany A.G., Bross P., Nandy A., Rasched I.,
Ghisla S.;
RT "Biochemical characterization of a variant human medium-chain acyl-CoA
dehydrogenase with a disease-associated mutation localized in the
active site.";
RL Biochem. J. 337:225-230(1999).
RN [17]
RC VARIANT HIS-67.
RA PubMed-11349232;
Andresen B.S., Dobrowolski S.F., O'Reilly L., Muenzer J.,
McCandless S.E., Frazier D.M., Udvari S., Bross P., Knudsen I.,
Banas R., Chace D.H., Engel P.C., Naylor E.W., Gregersen N.;
RT "Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by
MS/MS-based prospective screening of newborns differ from those
observed in patients with clinical symptoms: identification and
characterization of a new, prevalent mutation that results in mild
MCAD deficiency.";
RL Am. J. Hum. Genet. 68:1408-1418(2001).
RN [18]
RC VARIANT MCAD LEU-245.
RA PubMed-11409868;
Zschocke J., Schulze A., Lindner M., Fiesel S., Olgemoller K.,
Hoffmann G.F., Penzien J., Ruiter J.P.N., Wanders R.J.A.,
Mayatepek E.;

RT Molecular and functional characterization of mild MCAD deficiency.";
RL Hum. Genet. 108:404-408(2001).
RN [19]
RP VARIANT THR-281.
RX PubMed=11486912;
RA Albers S., Levy H.L., Irons M., Strauss A.W., Marsden D.;
RT "Compound heterozygosity in four asymptomatic siblings with
medium-chain acyl-CoA dehydrogenase deficiency.";
RL J. Inher. Metab. Dis. 24:417-418(2001).
CC -!- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO
16.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
ETF.
CC -!- COFACTOR: FAD.
CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
step.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- DISEASE: DEFECTS IN ACAD ARE THE CAUSE OF MCAD DEFICIENCY, AN
AUTOSOMAL RECESSIVE DISEASE WHICH CAUSES FASTING HYPOLYCEMIA,
HEPATIC DYSFUNCTION, AND ENCEPHALOPATHY, OFTEN RESULTING IN DEATH
IN INFANCY. THE DISEASE FREQUENCY IS ONE IN 13000.
CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-CoA DEHYDROGENASES
OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
TISSUES.
CC -!- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M91432; AAA59567.1; -;
DR EMBL; M91421; AAA59567.1; JOINED.
DR EMBL; M91422; AAA59567.1; JOINED.
Query Match 62.0%; Score 31; DB 1; Length 421;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGXDY 10
Db 58 EEIIPVAAEY 67
ID ACADM_PIG STANDARD; PRT; 421 AA.
AC P41367;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
(EC 1.3.99.3) (MCAD).
GN ACADM.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki H., Kimura M., Ito T., Murakami Y., Hamasima N., Yasue H.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93361479; PubMed=8356049;
RA Kim J.-P., Wang M., Paschke R.;
RT "Crystal structures of medium-chain acyl-CoA dehydrogenase from pig
liver mitochondria with and without substrate.";

Proc. Natl. Acad. Sci. U.S.A. 90:7523-7527(1993).
-!- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO
16.
-!- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
ETF.
-!- COFACTOR: FAD.
-!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
step.
-!- SUBUNIT: HOMOTETRAMER.
-!- SUBCELLULAR LOCATION: Mitochondrial matrix.
-!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-CoA DEHYDROGENASES
OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
TISSUES.
-!- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40845; AAA83759.1; -;
DR PDB; 3MDD; 30-SEP-94.
DR PDB; 3MDE; 30-SEP-94.
DR InterPro; IPR001552; Acyl-CoA_dh.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh.M; 1.
DR Pfam; PF02771; Acyl-CoA_dh.N; 1.
DR PROSITE; PS00072; ACYL-CoA_DH_1; 1.
DR PROSITE; PS00073; ACYL-CoA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW Mitochondrion; Transit peptide; 3D-structure.
FT TRANSIT 1 25 MITOCHONDRION.
FT CHAIN 26 421 ACYL-CoA DEHYDROGENASE, MEDIUM-CHAIN
SPECIFIC.
FT ACT_SITE 193 193 FORMS A HYDROGEN-BOND WITH THE FLAVIN
N(5) OF THE FAD COFACTOR (BY SIMILARITY).
FT ACT_SITE 401 401 BASE.
FT -CONFLICT 40 40 K -> E (IN REF. 2).
FT -CONFLICT 283 283 S -> P (IN REF. 2).
FT -CONFLICT 305 305 G -> E (IN REF. 2).
FT -CONFLICT 331 331 D -> E (IN REF. 2).
SQ SEQUENCE 421 AA; 46358 MW; 27347B0C99F2B7C3 CRC64;
Query Match 62.0%; Score 31; DB 1; Length 421;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGXDY 10
Db 58 EEIIPVAAEY 67
ID TR2M_ERWHE STANDARD; PRT; 562 AA.
AC Q47861;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Trypophan 2-monooxygenase (EC 1.13.12.3).
GN IAAM.
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=py. gypsophilae / PD713;
RA Ophir Y., Kunik T., Manulis S., Lichter A., Barash I., Gafni Y.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

QY 1 EEVWPXG 7
 DB 28 EEVWPGG 34

RESULT 22
 HES3_RAT
 ID HES3_RAT STANDARD; PRT; 175 AA.
 AC Q04667;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Transcription factor HES-3 (Hairy and enhancer of split 3).
 GN HES3 OR HES-3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=94040724; PubMed=1340473;
 RA Sasai Y., Kageyama R., Tagawa Y., Shigemoto R., Nakanishi S.;
 RT "Two mammalian helix-loop-helix factors structurally related to
 RT Drosophila hairy and enhancer of split.";
 RL Genes Dev. 6:2620-2634(1992).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
 CC PROTEIN FOR THEIR TRANSCRIPTION.
 CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
 CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN PURKINJE CELLS.
 CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
 CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
 CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
 CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 CC HAIRY-RELATED PROTEINS.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D13418; BAA02683.1; -;
 DR PIR: S36749; S36749.
 DR TRANSFAC: T01651; -;
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR PROSITE: PS00038; HLH_1; 1.
 DR PROSITE: PS00888; HLH_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Repressor.
 FT DNA_BIND 1 6
 FT DOMAIN 7 50 BASIC DOMAIN.
 FT DOMAIN 108 167 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 172 175 PRO-RICH.
 FT WRPW MOTIF (REQUIRED FOR ACTIVITY)
 FT (BY SIMILARITY).
 SQ SEQUENCE 175 AA; 19137 MW; D490663869155CB3 CRC64;

Query Match 60.0%; Score 30; DB 1; Length 175;
 Best Local Similarity 62.5%; Pred. NO. 30;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VWPXGXDY 10
 DB 58 LVPXGVDY 65

RESULT 23
 CTC_BACSU
 ID CTC_BACSU STANDARD; PRT; 203 AA.
 AC P14194;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE General stress protein ctc.
 GN CTC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96051385; PubMed=7584024;
 RA Ogasawara N., Nakai S., Yoshikawa H.;
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus
 RT subtilis chromosome containing the replication origin.";
 RL DNA Res. 1:1-14(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Friz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.F., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE OF 1-185 FROM N.A.
 RX MEDLINE=90066361; PubMed=2555671;
 RA Nilsson D., Hove-Jensen B., Arnvig K.;
 RT "Primary structure of the tms and prs genes of Bacillus subtilis.";
 RL Mol. Gen. Genet. 218:565-571(1989).
 RN [4]
 RP SEQUENCE OF 1-13.
 RC STRAIN=168 / IS58;
 RX Voelker U., Engelmann S., Maul B., Riethdorf S., Voelker A.,
 RA Schmid R., Mach H., Hecker M.;
 RT "Analysis of the induction of general stress proteins of Bacillus
 RT subtilis.";
 RL Microbiology 140:741-752(1994).
 CC -1- FUNCTION: NOT KNOWN.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT THE END OF EXPONENTIAL GROWTH
 CC UNDER CONDITIONS IN WHICH THE ENZYMES OF THE TCA CYCLE ARE
 CC REPRESSED.
 CC -!- INDUCTION: BY HEAT SHOCK, SALT STRESS, OXIDATIVE STRESS, GLUCOSE
 CC -!- LIMITATION: BY HEAT SHOCK, SALT STRESS, OXIDATIVE STRESS, GLUCOSE
 CC -!- SIMILARITY: BELONGS TO THE L25P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D26185; BAA05287.1; -
 CC EMBL; Z99104; CAB11828.1; -
 CC EMBL; X16518; CAA34524.1; -
 CC PIR; S05373; S05373.
 CC Subtilist; BG10115; etc.
 CC InterPro; IPR001021; Ribosomal_L25.
 CC Pfam; PF01386; Ribosomal_L25p; 1.
 CC ProDom; PD012503; Ribosomal_L25; 1.
 CC TIGRFAMs; TIGR00731; ctc.TL5; 1.
 CC Heat shock; Complete proteome.
 CC INIT MET 0
 CC VARIAT 1 1 A -> R (IN STRAIN IS58).
 CC VARIAT 9 9 T -> Q (IN STRAIN IS58).
 CC VARIAT 12 12 T -> I (IN STRAIN IS58).
 CC SEQUENCE 203 AA; 21924 MW; 5130A404D0B6247 CRC64;
 CC
 CC Query Match 60.0%; Score 30; DB 1; Length 203;
 CC Best Local Similarity 62.5%; Pred. No. 35;
 CC Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 4 VPXGDYXS 11
 CC : | | | |
 CC Db 157 LPAGGDYS 164
 CC
 CC RESULT 24
 CC ARAD_ECOLI STANDARD; PRT; 231 AA.
 CC ID ARAD_ECOLI STANDARD; PRT; 231 AA.
 CC AC P08203;
 CC DT 01-AUG-1988 (Rel. 08, Created)
 CC DT 01-AUG-1991 (Rel. 19, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE L-ribulose 5-phosphate 4-epimerase (EC 5.1.3.4) (Phosphoribulose
 CC isomerase).
 CC GN ARAD OR B0061.
 CC OS Escherichia coli.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Escherichia.
 CC OX NCBI_TaxID:562;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=B;
 CC RX MEDLINE=87163495; PubMed=3549454;
 CC RA Lee N., Gielow W., Martin R., Hamilton E., Fowler A.;
 CC RT "The organization of the arabad operon of Escherichia coli.";
 CC RL Gene 47:231-244(1986).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=K12;
 CC RX MEDLINE=91083835; PubMed=2261080;
 CC RA Chen H., Sun Y., Stark T., Beattie W., Moses R.E.;
 CC RT "Nucleotide sequence and deletion analysis of the polb gene of
 CC Escherichia coli.";
 CC RL DNA Cell Biol. 9:631-635(1990).
 CC RN [3]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=K12;
 CC RX MEDLINE=91067495; PubMed=2251150;

RA Mineno J., Fukui H., Ishino Y., Kato I., Shinagawa H.;
 RT "Nucleotide sequence of the arad gene of Escherichia coli K12
 RT encoding the L-ribulose 5-phosphate 4-epimerase.";
 RN Nucleic Acids Res. 18:6722-6722(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [6]
 RP SEQUENCE OF 222-231 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91017565; PubMed=2217198;
 RA Bonner C.A., Hays S., McEntee K., Goodman M.F.;
 RT "DNA polymerase II is encoded by the DNA damage-inducible dinA gene
 RT of Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:7663-7667(1990).
 RN [7]
 RP SEQUENCE OF 158-231 FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=91238699; PubMed=2034216;
 RA Iwasaki H., Ishino Y., Toh H., Nakata A., Shinagawa H.;
 RT "Escherichia coli DNA polymerase II is homologous to alpha-like DNA
 RT polymerases.";
 RL Mol. Gen. Genet. 226:24-33(1991).
 CC -!- CATALYTIC ACTIVITY: L-ribulose 5-phosphate = D-xylulose 5-
 CC phosphate.
 CC -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
 CC -!- PATHWAY: L-arabinose catabolism; third step.
 CC -!- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ARAD/FUCA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15263; AAA23464.1; -
 CC EMBL; M35371; -; NOT_ANNOTATED_CDS.
 CC EMBL; M62646; AAA24405.1; -
 CC EMBL; D10483; BAA01332.1; -
 CC EMBL; AE000116; AAC73172.1; -
 CC EMBL; M37727; AAA23683.1; -
 CC EMBL; M38283; AAA63763.1; -
 CC EMBL; X56048; CAA39519.1; -
 CC PIR; D29022; ISECP4.
 CC PIR; S13593; ISECK4.
 CC PIR; A36236; A36236.
 CC PIR; S40577; S40577.
 CC ECO2DBASE; G028.1; 6TH EDITION.
 CC EcoGene; Egl0055; arad.
 CC InterPro; IPR001303; Aldolase_II_N.
 CC InterPro; IPR004661; Arab.
 CC Pfam; PF00596; Aldolase_II; 1.
 CC TIGRFAMs; TIGR00760; arad; 1.
 CC Arabinose catabolism; Isomerase; Zinc; Complete proteome.
 KW

```
FT METAL          76 76 ZINC (BY SIMILARITY).
FT METAL          95 95 ZINC (BY SIMILARITY).
FT METAL          97 97 ZINC (BY SIMILARITY).
FT METAL          171 171 ZINC (BY SIMILARITY).
FT VARIANT        50 50 V -> I.
FT VARIANT        70 70 T -> A.
FT VARIANT        216 216 D -> N.
SQ SEQUENCE 231 AA; 25519 MW; 1753F75958332163 CRC64;

Query Match          60.0%; Score 30; DB 1; Length 231;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
   : | | | | |
DB 41 IKPSGVDYS 49

RESULT 25
ARAD_SALTY
ID ARAD_SALTY STANDARD; PRT; 231 AA.
AC P06190;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) (Phosphoribulose
isomerase).
GN ARAD OR STM0101.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=85232046; PubMed=3891514;
RA Lin H.-C., Lei S.-P., Studnicka G., Wilcox G.;
RT "The arabAD operon of Salmonella typhimurium LT2. III. Nucleotide
sequence of arabD and its flanking regions, and primary structure of
its product, L-ribulose-5-phosphate 4-epimerase."
RL Gene 34:129-134(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SCGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856(2001).
CC -!- CATALYTIC ACTIVITY: L-ribulose 5-phosphate = D-xylulose 5-
phosphate.
CC -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (POTENTIAL).
CC -!- PATHWAY: L-arabinose catabolism; third step.
CC -!- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ARAD/FUCA
SUBFAMILY.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
frameshift in position 202.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; M11047; AAA27025.1; ALT_FRAME.
CC EMBL; AE008698; AAL19065.1; -.
CC PIR; A24986; ISEB47.
```

```
DR StyGene: SG10015; arad.
DR InterPro; IPR001303; Aldolase_II_N.
DR InterPro; IPR004661; Arad.
DR Pfam; PF00596; Aldolase_II; 1.
DR TIGRFAMs; TIGR00760; arad; 1.
KW Arabinose catabolism; Isomerase; Zinc; Complete proteome.
FT METAL          76 76 ZINC (BY SIMILARITY).
FT METAL          95 95 ZINC (BY SIMILARITY).
FT METAL          97 97 ZINC (BY SIMILARITY).
FT METAL          171 171 ZINC (BY SIMILARITY).
SQ SEQUENCE 231 AA; 25531 MW; DA473505739284F9 CRC64;

Query Match          60.0%; Score 30; DB 1; Length 231;
Best Local Similarity 55.6%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPXGXDYS 11
   : | | | | |
DB 41 IKPSGVDYS 49
```

Search completed: June 10, 2003, 13:40:25
Job time : 5.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-46
Perfect score: 50
Sequence: 1 EEVVPXGDYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 75 summaries

Database : SPTREMBL_21.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organella.*
9: sp.phase.*
10: sp.plant.*
11: sp.podent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.rvirus.*
16: sp.bacteriaph.*
17: sp.bacteriaph.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	72.0	298	2	O52367 rhizobium t
2	36	72.0	363	17	O30260 archaeoglob
3	35	70.0	78	6	Q9xst4 canis famli
4	35	70.0	143	5	Q9vsv8 drosophila
5	35	70.0	217	4	O00404 homo sapien
6	35	70.0	290	16	Q8u7j0 agrobacteri
7	35	70.0	299	4	Q9uee9 homo sapien
8	35	70.0	587	16	Q9jzps neisseria m
9	35	70.0	692	5	Q9vi90 drosophila
10	35	70.0	906	10	Q9zsy4 arabidopsis
11	35	70.0	908	10	Q9fjk8 arabidopsis
12	35	70.0	908	10	Q8w4j9 arabidopsis
13	35	70.0	908	10	Q9zsy3 arabidopsis
14	35	70.0	909	10	Q9m5ai arabidopsis
15	35	70.0	1063	16	Q8rg86 fusobacteri
16	34	68.0	156	3	Q12479 saccharomyc

17	34	68.0	175	10	Q8vy88 arabidopsis
18	34	68.0	247	16	Q92u66 rhizobium m
19	34	68.0	257	10	Q9c6j0 arabidopsis
20	34	68.0	433	16	Q9a382 caulobacter
21	34	68.0	1442	17	Q96yh5 sulfolobus
22	34	68.0	2778	5	Q9v9t6 drosophila
23	33	66.0	143	17	Q8tx62 methanopyru
24	33	66.0	165	16	Q92bc5 listeria in
25	33	66.0	165	16	Q8v6u8 listeria mo
26	33	66.0	196	16	Q8zaq9 versinia pe
27	33	66.0	210	10	Q65890 cyclotella
28	33	66.0	225	10	Q40129 lycopersico
29	33	66.0	253	16	Q8xpa8 clostridium
30	33	66.0	312	11	Q9d876 mus musculu
31	33	66.0	440	17	Q9yfi3 aeropyrum p
32	33	66.0	471	11	Q8rl26 mus musculu
33	33	66.0	484	11	Q8vd18 mus musculu
34	33	66.0	517	16	Q8xz15 raietonia s
35	33	66.0	563	13	Q9ddj4 halichoeres
36	33	66.0	563	16	Q99xl5 streptococc
37	33	66.0	678	12	Q9elx6 cercopitheci
38	33	66.0	1028	16	Q8yjl1 bruceella me
39	33	66.0	1088	16	Q9kt32 vibrio chol
40	33	66.0	1150	5	O17704 caenorhabdi
41	33	66.0	1828	16	Q98k29 rhizobium l
42	32	64.0	105	11	Q9crg3 mus musculu
43	32	64.0	140	17	Q8tng3 methanosarc
44	32	64.0	154	10	Q9sbb8 oryza sativ
45	32	64.0	219	5	Q9gq04 eriocheir s
46	32	64.0	253	15	P88362 human immun
47	32	64.0	266	11	Q9jma5 mus musculu
48	32	64.0	275	5	Q9vix6 drosophila
49	32	64.0	280	5	Q9vnm9 drosophila
50	32	64.0	295	11	O70565 mus musculu
51	32	64.0	295	11	O88271 mus musculu
52	32	64.0	301	16	Q99td4 staphylococ
53	32	64.0	307	17	Q9hp60 halobacteri
54	32	64.0	314	16	Q8yyj8 anabaena sp
55	32	64.0	336	11	Q9d8m6 mus musculu
56	32	64.0	336	11	Q9cx34 mus musculu
57	32	64.0	354	11	Q9cre7 mus musculu
58	32	64.0	357	17	O29920 archaeoglob
59	32	64.0	366	17	O29451 archaeoglob
60	32	64.0	405	17	Q8ty23 methanopyru
61	32	64.0	423	16	Q92m00 rhizobium m
62	32	64.0	425	5	Q9xvk4 caenorhabdi
63	32	64.0	511	2	O52680 escherichia
64	32	64.0	543	3	Q8tff4 trichoderma
65	32	64.0	564	17	Q8u2a4 pyrococcus
66	32	64.0	565	16	Q9cin1 lactococcus
67	32	64.0	632	17	Q9hnp8 halobacteri
68	32	64.0	672	5	O45063 caenorhabdi
69	32	64.0	745	5	Q95p46 carcinus ma
70	32	64.0	873	10	Q9lkl4 zea mays (m
71	32	64.0	884	10	Q9af60 dicanthell
72	32	64.0	1031	5	Q9u6a3 callinectes
73	32	64.0	1082	10	Q94dz5 oryza sativ
74	32	64.0	1410	2	O52673 escherichia
75	32	64.0	1420	2	O52666 escherichia

ALIGNMENTS

RESULT 1
O52367
ID O52367; PRELIMINARY; PRT; 298 AA.
AC O52367; (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Aryl-alcohol dehydrogenase homolog (Fragment).
GN XYL81.

OS Rhizobium tropici.
 OG Plasmid pRtCFN299a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=398;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CFN299;
 RA Rosenblueth M., Hynes M.F., Martinez-Romero E.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -I- COFACTOR: ZINC (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY
 DR EMBL; AF036920; AAC04779.1; -
 DR HSP; P07846; ISDG.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; Adh_zn_family.
 DR Pfam; PF00107; adh_zinc; 1.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Plasmid; Zinc.
 FT NON_TER 298
 SQ SEQUENCE 298 AA; 31092 MW; 49B2F8117C33AE87 CRC64;
 Query Match 72.0%; Score 36; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPGXGXDYS 11
 Db [1] [1] [1] [1]
 250 EIIPEGADFS 259
 RESULT 2
 O30260 PRELIMINARY; PRT; 363 AA.
 ID O30260
 AC O30260
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein AF2411.
 GN AF2411.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Richardson D.L., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Fleischmann R.D., Kerlavage A.R., Graham D.E., Kyrtides N.C.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.J., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91255.1; -
 DR TIGR; AF2411;
 DR InterPro; IPR002103; Bac_luciferase.
 DR Pfam; PF00296; bac_luciferase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 363 AA; 41736 MW; 0E976EAE78F4803 CRC64;
 Query Match 72.0%; Score 36; DB 17; Length 363;
 Best Local Similarity 54.5%; Pred. No. 15;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXGXDYS 11
 Db [1] [1] [1] [1]
 120 ENIVPYGIDFS 130
 RESULT 3
 Q9XST4 PRELIMINARY; PRT; 78 AA.
 ID Q9XST4
 AC Q9XST4
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE p97 homologous protein (Fragment).
 GN P97.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYROID;
 RX MEDLINE=20422104; PubMed=10964405;
 RA Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C.,
 RA Christophe D.;
 RT "A method for the large-scale cloning of nuclear proteins and nuclear
 RT targeting sequences on a functional basis.";
 RL Anal. Biochem. 284:231-239(2000).
 DR EMBL; AJ388531; CAB46833.1; -
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8895 MW; B62486313555FBA1 CRC64;
 Query Match 70.0%; Score 35; DB 6; Length 78;
 Best Local Similarity 54.5%; Pred. No. 4.4;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGXGXDYS 11
 Db [1] [1] [1] [1]
 16 EDVPSGGEYS 26
 RESULT 4
 Q9VSY8 PRELIMINARY; PRT; 143 AA.
 ID Q9VSY8
 AC Q9VSY8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG3911 protein.
 GN CG3911.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Ballif J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabriellian A.E., Gaig N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003552; AAF50270.1; -.
 DR FlyBase; FBgn0035992; CG3911.
 SQ SEQUENCE 143 AA; 16471 MW; 1DC346DC22C02AA2 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 143;
 Best Local Similarity 66.7%; Pred. No. 8.7;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXD 9
 Db 135 EEVVPAGED 143

RESULT 5

ID 000404 PRELIMINARY; PRT; 217 AA.

AC 000404;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE P97 homologous protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=97160586; PubMed=9006920;

RA Nobukuni T., Kobayashi M., Omori A., Ichinose S., Iwanaga T.,
 RA Takahashi I., Hashimoto K., Hattori S., Kaibuchi K., Miyata Y.,
 RT "An Alu-linked repetitive sequence corresponding to 280 amino acids is
 expressed in a novel bovine protein, but not in its human homologue."
 RL J. Biol. Chem. 272:2801-2807(1997).
 DR EMBL; D85939; BAA20069.1; -.
 SQ SEQUENCE 217 AA; 24061 MW; E404BA2E35497828 CRC64;

Query Match 70.0%; Score 35; DB 4; Length 217;
 Best Local Similarity 54.3%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDS 11
 Db 16 EDVVPSGEYS 26

RESULT 6

Q8U7J0

ID Q8U7J0 PRELIMINARY; PRT; 290 AA.
 AC Q8U7J0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 6-O-methylguanine-DNA methyltransferase.
 GN ADA OR ATU4459 OR AGR_L_818.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58."
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Roumel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009374; AAL45253.1; -.
 DR EMBL; AE008240; AAK8982.1; -.
 KW Methyltransferase; Transference; Complete proteome.
 SQ SEQUENCE 290 AA; 31587 MW; B626592EF519977F CRC64;

Query Match 70.0%; Score 35; DB 16; Length 290;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDX 10
 Db 9 EDITPIGSDY 18

RESULT 7
 Q8U7J0 PRELIMINARY; PRT; 299 AA.
 ID Q8U7J0
 AC Q8U7J0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BCNT protein (CRANIOFACIAL development protein 1).
 OS BCNT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98267221; PubMed=9602175;
 RA Takahashi I., Nobukuni T., Omori H., Kobayashi M., Tanaka S.,
 RA Ohshima K., Okada N., Masui T., Hashimoto K., Iwashita S.;
 RT "Existence of a bovine LINE repetitive insert that appears in the CDNA
 of bovine protein BCNT in ruminant, but not in human, genomes.";

```

RL  Gene 211:387-394(1998).
RN  [2]
RC  SEQUENCE FROM N.A.
RP  TISSUE=PLACENTA;
RA  Strausberg R.;
RL  Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB09285; BAA31867.1; -
DR  EMBL; BC000991; AAH00991.1; -
SQ  SEQUENCE 299 AA; 33593 MW;  F4A9E928B669451A CRC64;

Query Match          70.0%; Score 35; DB 4; Length 299;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEVVPXGXDYS 11
Db  16 EDYVPSGGEYS 26

RESULT 8
Q9JZP8 PRELIMINARY; PRT; 587 AA.
AC Q9JZP8:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE succinate dehydrogenase, flavoprotein subunit.
GN NMB0950.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittoni H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT *Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.
RL Science 287:1809-1815(2000).
CC -1- COFACTOR: FAD (BY SIMILARITY).
DR EMBL; AE002446; AAF41356.1; -
DR HSSP; P00363; 1FUM.
DR TIGR; NMB0950; -
DR InterPro; IPR003953; FAD_bind2.
DR InterPro; IPR001327; FAD_Pyr_redox.
DR InterPro; IPR003952; FRD/SDH_FAD.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR004112; Succ_DH_flav_C.
DR Pfam; PF00890; FAD_binding_2; 1.
DR Pfam; PF02910; succ_DH_flav_C; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNRDRTASEI.
DR PROSITE; PS00504; FRD_SDH_FAD_BINDING; 1.
DR PROSITE; PS00504; Oxidoreductase; Complete proteome.
KW FAD; Flavoprotein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 587 AA; 64502 MW; 9581701B08069003 CRC64;

Query Match          70.0%; Score 35; DB 16; Length 587;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEVVPXGXDY 10
Db  366 EVVVPQGEDY 375

RESULT 9
Q9VI90 PRELIMINARY; PRT; 692 AA.
AC Q9VI90:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG10040 protein (Roughened eye).
GN RN OR CG10040 OR CG14600 OR CG14601.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC St Pierre S.E., Galindo M.I., Couso J.P., Thor S.;
RT "Control of Drosophila imaginal disc development by rotund and
RT roughened eye: differentially expressed transcripts of the same gene
RT encoding distinct zinc finger proteins."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003672; AAF54035.1; -
DR EMBL; AF395904; AAL59598.1; -
DR HSSP; P08153; 1ZFD.
DR FlyBase; FBgn0037494; rn.
DR InterPro; IPR000822; znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR ProDom; PD000003; znf_C2H2; 1.
DR SMART; SM00355; znf_C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
KW DNA-binding; Metal-binding; zinc-finger.

```

SQ SEQUENCE 692 AA; 75774 MW; BD010502BB65042E CRC64;

Query Match 70.0%; Score 35; DB 5; Length 692;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVVPXGXDYS 11
 I:| | | | |
 Db 26 EIPPAGGDYS 35

RESULT 10

Q9ZSY4 PRELIMINARY; PRT; 906 AA.

AC Q9ZSY4;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein RPP8.
 GN RPP8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LA-ER;
 RX MEDLINE=99030193; PubMed=9811794;
 RA McDowell J.M., Dhandaaydham M., Long T.A., Aarts M.G., Goff S.,
 RA Holub E.B., Dangl J.L.;
 RT "Intragenic recombination and diversifying selection contribute to the
 RT evolution of downy mildew resistance at the RPP8 locus of
 RT Arabidopsis.";
 RL Plant Cell 10:1861-1874 (1998).
 DR EMBL; AF089710; AAC83165.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 1.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 906 AA; 104201 MW; 52905EB143676F8F CRC64;

Query Match 70.0%; Score 35; DB 10; Length 906;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 I:| | | | |
 Db 881 EKLVPGGEDY 890

RESULT 11

Q9FJK8 PRELIMINARY; PRT; 908 AA.

AC Q9FJK8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99087489; PubMed=9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.

RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:297-308 (1998).
 DR EMBL; AB015468; BAB10695.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104448 MW; 3111991B17239693 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 I:| | | | |
 Db 883 EKLVPGGEDY 892

RESULT 12

Q8W4J9 PRELIMINARY; PRT; 908 AA.

AC Q8W4J9;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Disease resistance protein RPP8.
 GN AT5G43470, MW20.19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY062514; AAL32592.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104649 MW; 4461F553128F3A15 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGXDY 10
 I:| | | | |
 Db 883 EKLVPGGEDY 892

RESULT 13

Q9ZSY3 PRELIMINARY; PRT; 908 AA.

AC Q9ZSY3;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RPP8 (Disease resistance protein RPP8).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COL;
 RX MEDLINE=99030193; PubMed=9811794;
 RA McDowell J.M., Dhondydam M., Long T.A., Aarts M.G.M., Goff S.,
 RA Holub E.B., Dangl J.L.;
 RT "Intragenic recombination and diversifying selection contribute to the
 RT evolution of downy mildew resistance at the RPP8 locus of
 RT Arabidopsis";
 RL Plant Cell 10:1861-1874(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty p1 and TAC
 RT clones";
 RL DNA Res. 7:31-63(2000).
 DR EMBL; AF089711; AAC78631.1; -;
 DR EMBL; AB025638; BAA97426.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 908 AA; 104681 MW; 1BFA35BB6B0CB5CD CRC64;
 Query Match 70.0%; Score 35; DB 10; Length 908;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 Db 883 EKLVPGGEDY 892
 I::||| |
 RESULT 14
 Q9M5A1
 ID Q9M5A1 PRELIMINARY; PRT; 909 AA.
 AC Q9M5A1
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Viral resistance protein.
 GN HRT.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DI-17;
 RA Cooley M.B., Pathirana S., Wu H., Kachroo P., Klessig D.F.;
 RT "Members of the Arabidopsis HRT/RPP8 family of resistance genes confer
 RT resistance to both viral and oomycete pathogens.";
 RL Plant Cell 0:0-0(2000).
 DR EMBL; AF234174; AAR36987.1; -;
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 SQ SEQUENCE 909 AA; 105052 MW; 06262B71A2B3037F CRC64;
 Query Match 70.0%; Score 35; DB 10; Length 909;
 Best Local Similarity 60.0%; Pred. No. 70;
 QY 1 EEVVPXGXDY 10
 Db 883 EKLVPGGEDY 892
 I::||| |
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EEVVPXGXDY 10
 Db 884 EKLVPGGEDY 893
 I::||| |
 RESULT 15
 Q8RG86
 ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
 AC Q8RG86;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
 GN FN0422
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteriia; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010554; AAL94625.1; -;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCB411 CRC64;
 Query Match 70.0%; Score 35; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 84;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 EVVPXGXDYS 11
 Db 195 EIVPGLNYS 204
 I::|| |
 RESULT 16
 Q12479
 ID Q12479 PRELIMINARY; PRT; 156 AA.
 AC Q12479;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ORF YOR013W.
 GN YOR013W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC De haan M., Grivell L.A., Maarse A.C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FY1679;
 RA De haan M., Maarse A.C., Grivell L.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FY1679;

```

RX MEDLINE=94019318; PubMed=8413243;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RA Sherman F.;
RT "CYC2 encodes a factor involved in mitochondrial import of yeast
RT cytochrome c.";
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
DR EMBL; Z74920; CAA99201.1; -
DR EMBL; X87331; CAA60762.1; -
DR SGD; S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 68.0%; Score 34; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVXPXGXDY 10
DQ 50 EVVPLGMDY 58

RESULT 17
Q8VY88 PRELIMINARY; PRT; 175 AA.
AC Q8VY88;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 19.2 kDa protein.
GN ATIC50910.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072348; AAL61955.1; -
DR InterPro; IPR002110; ANK.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 175 AA; 19189 MW; 7F72AB1EC82C4190 CRC64;

Query Match 68.0%; Score 34; DB 10; Length 175;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGXDS 11
DQ 135 EELLKAGADYS 145

RESULT 18
Q92U66 PRELIMINARY; PRT; 247 AA.
AC Q92U66;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

```

```

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein RBL273.
GN RBL273 OR SMB21444.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603646; CAC49673.1; -
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 247 AA; 28930 MW; 2F14F383E66D420 CRC64;

Query Match 68.0%; Score 34; DB 16; Length 247;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGXDX 10
DQ 48 EDVEPRGADY 57

RESULT 19
Q9C6J0 PRELIMINARY; PRT; 257 AA.
AC Q9C6J0;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 28.6 kDa protein.
GN F8A12.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dewar K.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cresasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Miltitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC079284; RAG50930.1; -
DR InterPro; IPR002110; ANK.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat; Hypothetical protein; Repeat.

```

SQ SEQUENCE 257 AA; 28578 MW; 714C7A4387F32B5C CRC64;
 Query Match 68.0%; Score 34; DB 10; Length 257;
 Best Local Similarity 54.5%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDXYS 11
 DB 217 EELLKAGADYS 227

 RESULT 20
 Q9A382 PRELIMINARY; PRT; 433 AA.
 AC Q9A382;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Peptidoglycan-binding protein, putative.
 GN CC3322
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Eisen J., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Potocka I., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smt J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback J., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005994; AAK25284.1;
 DR HSSP; P41052; LITM.
 DR TIGR; CC3322;
 DR InterPro; IPR002477; PG_binding.
 DR Pfam; PF01471; PG_binding_1;
 KW Complete proteome.
 SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

 Query Match 68.0%; Score 34; DB 16; Length 433;
 Best Local Similarity 54.5%; Pred. No. 51;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDXYS 11
 DB 266 EVILPPGFDYS 276

 RESULT 21
 Q96YH5 PRELIMINARY; PRT; 1442 AA.
 AC Q96YH5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein SP2195.
 GN SP2195.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kavarabiyasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,

RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 RL Crenarchaeon, Sulfolobus tokodaii strain7";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000989; BAB67302.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1442 AA; 15649 MW; D63EC2C35228121F CRC64;

Query Match 68.0%; Score 34; DB 17; Length 1442;
 Best Local Similarity 50.0%; Pred. No. 2e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 1 EEVVPXGXDX 10
 DB 863 EEITPGANY 872

RESULT 22

Q9V9T6 PRELIMINARY; PRT; 2778 AA.
 AC Q9V9T6; Q9Y0Z7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE FAF protein (BCDNA:LD22582)
 GN FAF OR BCDNA:LD22582 OR CCI945.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liang D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Ljung Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter G.M., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE OF 1090-2778 FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacled J.M.,
 RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
 RA Celniker S.E.;
 RT "Full length *Drosophila melanogaster* cDNA sequence.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003779; AAF57198.1; -;
 DR EMBL; AF145677; AAD38652.1; -;
 DR FlyBase; FBgn0005632; faf.
 DR InterPro; IPR001394; UCH-2.
 DR Pfam; PF00442; UCH-1; 1.
 DR Pfam; PF00443; UCH-2; 1.
 DR PROSITE; PS00972; UCH_2.1; 1.
 DR PROSITE; PS00973; UCH_2.2; 1.
 DR PROSITE; PS0235; UCH_2.3; 1.
 SQ SEQUENCE 2778 AA; 311140 MW; FFB90438BA53A02B CRC64;
 Query Match 68.0%; Score 34; DB 5; Length 2778;
 Best Local Similarity 54.5%; Pred. No. 4.le+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGXDXS 11
 I : I I I I I
 Db 1394 EVIIPDGDQS 1404
 RESULT 23
 Q8TX62 PRELIMINARY; PRT; 143 AA.
 ID Q8TX62
 AC Q8TX62;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Uncharacterized conserved protein.
 GN MK0814.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OC NCBI_TaxID=2320;
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE-21927647; PubMed-11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;
 RT "The complete genome of hyperthermophile *Methanopyrus kandleri* AV19
 RT and monophyly of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010372; AAO0207.1; -;
 KW Complete proteome.
 SQ SEQUENCE 143 AA; 15734 MW; 4CB828A1FBBDDB CRC64;

Query Match 66.0%; Score 33; DB 17; Length 143;
 Best Local Similarity 60.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGXDX 10
 I : I I I I I
 Db 75 EELVPQAGY 84
 RESULT 24
 Q92BC5

ID Q92BC5 PRELIMINARY; PRT; 165 AA.
 AC Q92BC5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein lin1625.
 GN LIN1625.
 OS *Listeria innocua*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OC NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed-11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madoeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
 RT "Comparative genomics of *Listeria* species.";
 RL Science 294:849-852(2001).
 DR EMBL; AL596169; CAC96856.1; -;
 DR ListList; LIN01625; -;
 DR InterPro; IPR000866; AhpC-TSA.
 DR InterPro; IPR02065; TPX.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; TPX; UNKNOWN_1.
 DR PROSITE; PS01265; TPX; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 18162 MW; 77705B7CD8BC8F4D CRC64;
 Query Match 66.0%; Score 33; DB 16; Length 165;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EVVPXGXDX 10
 I : I I I I I
 Db 144 EVVPEGSDH 152
 RESULT 25
 Q8Y608 PRELIMINARY; PRT; 165 AA.
 ID Q8Y608
 AC Q8Y608;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein lmo1583.
 GN LMO1583.
 OS *Listeria monocytogenes*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OC NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-EGD-E / SEROVAR 1/2A;
 RX MEDLINE-21537279; PubMed-11679669;
 RA Baquero F., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madoeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of *Listeria species*."
 RL Science 294:849-852(2001).
 DR EMBL; AL591979; CAC99661.1; -.
 DR ListiList; LMO01583; -.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR002065; TPX.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 165 AA; 18133 MW; 77705B7CC46D424D CRC64;

Query Match 66.0%; Score 33; DB 16; Length 165;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 EWVPXGXDY 10
 |||||
 Db 144 EVVPEGSDH 152

Search completed: June 10, 2003, 13:46:39
 Job time : 26.7857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds

(without alignments)

46.744 Million cell updates/sec

Title: US-09-909-164-47

Perfect score: 52

Sequence: 1 EGVVPXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database: A_Geneseq.101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	96.2	11	23	ABB80521 Hepatitis C virus
2	50	96.2	11	23	ABB80522 Hepatitis C virus
3	50	96.2	11	23	ABB80525 Hepatitis C virus
4	50	96.2	11	23	ABB80526 Hepatitis C virus
5	50	96.2	11	23	ABB80529 Hepatitis C virus
6	50	96.2	11	23	ABB80559 Hepatitis C virus
7	50	96.2	11	23	ABB80563 Hepatitis C virus
8	50	96.2	11	23	ABB80564 Hepatitis C virus
9	50	96.2	11	23	ABB80565 Hepatitis C virus
10	50	96.2	11	23	ABB80566 Hepatitis C virus

11	50	96.2	11	23	ABB80568 Hepatitis C virus
12	46	88.5	11	23	ABB80524 Hepatitis C virus
13	46	88.5	11	23	ABB80528 Hepatitis C virus
14	46	88.5	11	23	ABB80529 Hepatitis C virus
15	46	88.5	11	23	ABB80561 Hepatitis C virus
16	46	88.5	11	23	ABB80562 Hepatitis C virus
17	45	86.5	11	23	ABB80523 Hepatitis C virus
18	45	86.5	11	23	ABB80527 Hepatitis C virus
19	45	86.5	11	23	ABB80535 Hepatitis C virus
20	45	86.5	11	23	ABB80536 Hepatitis C virus
21	45	86.5	11	23	ABB80539 Hepatitis C virus
22	45	86.5	11	23	ABB80540 Hepatitis C virus
23	45	86.5	11	23	ABB80558 Hepatitis C virus
24	45	86.5	11	23	ABB80560 Hepatitis C virus
25	44	84.6	11	23	ABB80544 Hepatitis C virus
26	44	84.6	11	23	ABB80545 Hepatitis C virus
27	44	84.6	11	23	ABB80549 Hepatitis C virus
28	44	84.6	11	23	ABB80552 Hepatitis C virus
29	44	84.6	11	23	ABB80553 Hepatitis C virus
30	42	80.8	11	23	ABB80530 Hepatitis C virus
31	41	78.8	11	23	ABB80538 Hepatitis C virus
32	41	78.8	11	23	ABB80542 Hepatitis C virus
33	41	78.8	11	23	ABB80543 Hepatitis C virus
34	40	76.9	11	23	ABB80537 Hepatitis C virus
35	40	76.9	11	23	ABB80541 Hepatitis C virus
36	40	76.9	11	23	ABB80547 Hepatitis C virus
37	40	76.9	11	23	ABB80548 Hepatitis C virus
38	40	76.9	11	23	ABB80551 Hepatitis C virus
39	40	76.9	11	23	ABB80556 Hepatitis C virus
40	40	76.9	11	23	ABB80557 Hepatitis C virus
41	40	76.9	20	20	AAU76810 Novel human diago
42	40	76.9	1022	22	ABG03621 Novel human diago
43	40	76.9	1022	22	ABG03826 Hepatitis C virus
44	40	76.9	1022	22	ABG08173 Hepatitis C virus
45	39	75.0	11	23	ABB80546 Hepatitis C virus
46	39	75.0	11	23	ABB80550 Hepatitis C virus
47	39	75.0	11	23	ABB80554 Hepatitis C virus
48	39	75.0	11	23	ABB80555 Hepatitis C virus
49	38	73.1	11	23	ABB80533 Hepatitis C virus
50	38	73.1	11	23	ABB80534 Hepatitis C virus
51	38	73.1	3472	21	AA190913 Hepatitis C virus
52	37	71.2	11	23	ABB80531 Hepatitis C virus
53	37	71.2	11	23	ABB80532 Hepatitis C virus
54	36	69.2	244	21	AA12881 Murine JNK3 bindin
55	36	69.2	484	21	AA12882 Murine JNK3 bindin
56	35	67.3	11	18	AAW99288 Peptide N424 from
57	34	65.4	842	21	AA143359 P. chrysogenum sut
58	34	65.4	947	21	AA25105 Pinus radiata cell
59	34	65.4	1070	22	AAU14378 Human novel protei
60	33	63.5	12	21	AAV83772 HCV NS3 protease s
61	33	63.5	12	21	AAV83774 HCV NS3 protease s
62	33	63.5	13	18	AAW99276 Peptide D4 from WO
63	33	63.5	13	18	AAW99275 Peptide 5 used in
64	33	63.5	14	18	AAW99275 Peptide 4 used in
65	33	63.5	14	18	AAW99277 Peptide C0 from WO
66	33	63.5	14	18	AAW99275 Peptide 4 used in
67	33	63.5	14	18	AAW99277 Peptide 6 used in
68	33	63.5	16	18	AAW99274 Peptide D2 from WO
69	33	63.5	16	18	AAW99284 Peptide 3 used in
70	33	63.5	18	18	AAW99273 Peptide D1 from WO
71	33	63.5	18	18	AAW99273 Peptide 2 used in
72	33	63.5	18	19	AAW71281 Cleavable substrate
73	33	63.5	20	18	AAW99272 Peptide PS from WO
74	33	63.5	20	18	AAW99272 Peptide 1 used in
75	33	63.5	20	20	AA14511 HCV NS3 protease t

ALIGNMENTS

RESULT 1
ABB80521
ID ABB80521 standard; peptide; 11 AA.

```

XX AC ABB80521;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX XX Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 64; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 96.2%; Score 50; DB 23; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.0011;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11
RESULT 2
ABB80522
ID ABB80522 standard; peptide; 11 AA.
XX AC ABB80522;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.

```

```

OS Synthetic.
XX Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX XX WO200208251-A2.
XX PN 31-JAN-2002.
XX PD 19-JUL-2001; 2001WO-US23169.
XX PF 21-JUL-2000; 2000US-220101P.
XX PR (CORV-) CORVAS INT INC.
XX PA Lim-wilby M, Levy OE, Brunck TK;
XX PI WPI; 2002-361643/39.
XX XX Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 64; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 96.2%; Score 50; DB 23; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.0011;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11
RESULT 3
ABB80525
ID ABB80525 standard; peptide; 11 AA.
XX AC ABB80525;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX OS virucide.
XX XX Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"

```

FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX Claim 17; Page 64; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EEVVPXGMSYS 11
DB 1 EEVVPXGMSYS 11
RESULT 4
ABB80526
ID ABB80526 standard; peptide; 11 AA.
XX
AC ABB80526;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX

XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX Claim 17; Page 64; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EEVVPXGMSYS 11
DB 1 EEVVPXGMSYS 11
RESULT 5
ABB80559
ID ABB80559 standard; peptide; 11 AA.
XX
AC ABB80559;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 8 /note= "Oxymethionine"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX

XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX
 XX Sequence 11 AA;
 XX
 XX Query Match 96.2%; Score 50; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB |||||
 1 EEVVPXGMSYS 11
 RESULT 6
 ABB80563
 ID ABB80563 standard; peptide; 11 AA.
 XX AC ABB80563;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 XX Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with
 XX Modified-site 11 residue 7"
 XX /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -

XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.
 XX
 XX Sequence 11 AA;
 XX
 XX Query Match 96.2%; Score 50; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB |||||
 1 EEVVPXGMSYS 11
 RESULT 7
 ABB80564
 ID ABB80564 standard; peptide; 11 AA.
 XX AC ABB80564;
 XX 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 XX Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with
 XX Modified-site 11 residue 7"
 XX /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.

```

XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11

RESULT 8
ABB80565
ID ABB80565 standard; peptide; 11 AA.
AC ABB80565;
XX
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "Norleucyl carbonyl forming keto-amide linkage
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11

RESULT 9
ABB80566
ID ABB80566 standard; peptide; 11 AA.
XX
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyl carbonyl residue forming a
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11

RESULT 10
ABB80567
ID ABB80567 standard; peptide; 11 AA.
XX
XX 08-OCT-2002 (first entry)
XX
XX
XX
XX

```

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 XX Synthetic.
 OS
 PH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 6 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT Modified-site 6 /note= "(s,s)allothreonyl carbonyl residue forming a
 FT Modified-site 11 keto-amide linkage with residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 PF 21-JUL-2000; 2000US-220101P.
 PR (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 PS Sequence 11 AA;
 XX Query Match 96.2%; Score 50; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMSYS 11
 DE
 DE
 DE 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 PH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT Misc-difference 9 residue 7"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 PN
 FT

FT Modified-site 6 /note= "Alpha-propynyl-glycyl-carbonyl residue forming
 FT Modified-site 11 a keto-amide linkage with residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 PS Sequence 11 AA;
 XX Query Match 96.2%; Score 50; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMSYS 11
 DE
 DE
 DE 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 PH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT Misc-difference 9 residue 7"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX WO200208251-A2.
 PN
 FT

PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 88.5%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0075;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMDYS 11
 RESULT 13
 ABB80528
 ID ABB80528 standard; peptide; 11 AA.
 XX
 AC ABB80528;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN
 XX 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 88.5%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0075;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMDYS 11
 RESULT 14
 ABB80529
 ID ABB80529 standard; peptide; 11 AA.
 XX
 AC ABB80529;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 9 /note= "D-form residue"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN
 XX 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C

PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 88.5%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0075;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMDYS 11
 RESULT 14
 ABB80529
 ID ABB80529 standard; peptide; 11 AA.
 XX
 AC ABB80529;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 9 /note= "D-form residue"
 FT Modified-site 11 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN
 XX 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C

PT virus protease -
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 88.5%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0075;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 Db 1 EEVVPXGMDYS 11
 |||||
 RESULT 15
 ABB80561
 ID ABB80561 standard; peptide; 11 AA.
 XX
 AC ABB80561;
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 PR (CORV-) CORVAS INT INC.
 XX
 PA Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 PS The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 Query Match 88.5%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0075;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 Db 1 EEVVPXGMDYS 11
 |||||
 RESULT 16
 ABB80562
 ID ABB80562 standard; peptide; 11 AA.
 XX
 AC ABB80562;
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 PN 31-JAN-2002.
 XX
 PD 19-JUL-2001; 2001WO-US23169.
 XX
 PF 21-JUL-2000; 2000US-220101P.
 XX
 PR (CORV-) CORVAS INT INC.
 XX
 PA Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 PS The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

SQ Query Match 88.5%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0075;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEWVPXGMSYS 11
Db 1 EEWVPXGMDYS 11

RESULT 17
ABB80523
ID ABB80523 standard; peptide; 11 AA.

XX AC ABB80523;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

XX Query Match 86.5%; Score 45; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.012;

XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX RESULT 19

XX ABB80535

XX ID ABB80535 standard; peptide; 11 AA.

QY 1 EEWVPXGMSYS 11
Db 1 EEWVPXGMHYS 11

RESULT 18

ABB80527

ID ABB80527 standard; peptide; 11 AA.

XX AC ABB80527;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

XX Query Match 86.5%; Score 45; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.012;

XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX RESULT 19

XX ABB80535

XX ID ABB80535 standard; peptide; 11 AA.

XX ABB80535;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #15.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6
XX Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX residue 7"
XX /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX Claim 17; Page 64; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
XX Query Match 86.5%; Score 45; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.012;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX 1 EEVVPXGMSYS 11
XX 1 EEVVPXGQSYS 11
XX RESULT 20
XX ABB80536
XX ID ABB80536 standard; peptide; 11 AA.
XX AC ABB80536;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #16.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX

OS Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX residue 7"
XX Misc-difference 9 /note= "D-form residue"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease
XX Claim 17; Page 64; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
XX Query Match 86.5%; Score 45; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.012;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX 1 EEVVPXGMSYS 11
XX 1 EEVVPXGQSYS 11
XX RESULT 21
XX ABB80539
XX ID ABB80539 standard; peptide; 11 AA.
XX AC ABB80539;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX residue 7"
XX

FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ Query Match 86.5%; Score 45; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11
RESULT 22
ABB80540
ID ABB80540 standard; peptide; 11 AA.
XX AC ABB80540;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6
FT Modified-site 8 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 11 residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.

XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX Sequence 11 AA;
SQ Query Match 86.5%; Score 45; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EEVVPXGMSYS 11
Db 1 EEVVPXGMSYS 11
RESULT 23
ABB80558
ID ABB80558 standard; peptide; 11 AA.
XX AC ABB80558;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 8 residue 7"
FT Modified-site 11 /note= "Oxymethionine"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 86.5%; Score 45; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.012; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMHYS 11
 RESULT 24
 ABB80560
 ID ABB80560 standard; peptide; 11 AA.
 XX AC ABB80560;
 XX DT 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT Modified-site 6 residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 86.5%; Score 45; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.012; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 DB 1 EEVVPXGMHYS 11
 RESULT 25
 ABB80544
 ID ABB80544 standard; peptide; 11 AA.
 XX AC ABB80544;
 XX DT 08-OCT-2002 (first entry)
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 KW Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 65; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX
SQ Sequence 11 AA;

Query Match 84.6%; Score 44; DB 23; Length 11;
Best Local Similarity 90.9%; Pred No. 0.019;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EEVVPXGMSYS 11
| | | | | | | | | |
Db 1 EEVVPXGTSYS 11

Search completed: June 10, 2003, 13:39:12
Job time : 31.3571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds

(without alignments)
33.564 Million cell updates/sec

Title: US-09-909-164-47

Perfect score: 52

Sequence: 1 EEVVPKMGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	65.4	947	4	US-09-228-986-73
2	33	63.5	45	2	US-08-637-759B-236
3	33	63.5	45	3	US-08-871-355A-236
4	33	63.5	45	4	US-09-201-945-236
5	33	63.5	65	6	5177197-51
6	33	63.5	410	6	5177197-1
7	33	63.5	1394	6	5177197-30
8	32	61.5	10	4	US-09-357-952-66
9	32	61.5	10	4	US-09-521-650-66
10	32	61.5	10	4	US-09-168-888-66
11	32	61.5	102	2	US-08-580-988A-23
12	32	61.5	152	2	US-08-460-694-4
13	32	61.5	152	3	US-08-460-744-4
14	32	61.5	152	3	US-07-667-711B-4
15	32	61.5	173	1	US-08-193-977-7
16	32	61.5	189	2	US-08-464-517-21
17	32	61.5	189	2	US-08-246-361A-21
18	32	61.5	189	3	US-08-463-772-21
19	32	61.5	189	5	PCT-US93-05000-21
20	32	61.5	236	2	US-08-464-517-22
21	32	61.5	236	2	US-08-246-361A-22
22	32	61.5	236	3	US-08-463-772-22
23	32	61.5	236	5	PCT-US93-05000-22
24	32	61.5	280	2	US-08-464-517-6
25	32	61.5	280	3	US-08-463-772-6
26	32	61.5	289	2	US-08-246-361A-4
27	32	61.5	289	5	PCT-US93-05000-4

28	32	61.5	291	5	PCT-US93-05000-6	Sequence 6, Appli
29	32	61.5	292	2	US-08-464-517-23	Sequence 23, Appl
30	32	61.5	292	2	US-08-246-361A-6	Sequence 6, Appli
31	32	61.5	292	2	US-08-246-361A-23	Sequence 23, Appl
32	32	61.5	292	3	US-08-463-772-23	Sequence 23, Appl
33	32	61.5	292	5	PCT-US93-05000-23	Sequence 23, Appl
34	32	61.5	295	1	US-07-947-120-8	Sequence 8, Appli
35	32	61.5	295	1	US-08-472-893A-8	Sequence 8, Appli
36	32	61.5	295	2	US-08-460-694-2	Sequence 2, Appli
37	32	61.5	295	2	US-08-464-517-19	Sequence 19, Appl
38	32	61.5	295	2	US-08-464-517-20	Sequence 20, Appl
39	32	61.5	295	2	US-08-246-361A-19	Sequence 19, Appl
40	32	61.5	295	2	US-08-246-361A-20	Sequence 20, Appl
41	32	61.5	295	3	US-08-463-772-19	Sequence 19, Appl
42	32	61.5	295	3	US-08-463-772-20	Sequence 20, Appl
43	32	61.5	295	3	US-08-460-744-2	Sequence 2, Appli
44	32	61.5	295	3	US-07-667-711B-2	Sequence 2, Appli
45	32	61.5	295	3	US-08-947-492-8	Sequence 8, Appli
46	32	61.5	295	5	PCT-US93-05000-2	Sequence 2, Appli
47	32	61.5	295	5	PCT-US93-05000-19	Sequence 19, Appl
48	32	61.5	295	5	PCT-US93-05000-20	Sequence 20, Appl
49	32	61.5	309	2	US-08-464-517-4	Sequence 4, Appli
50	32	61.5	309	3	US-08-463-772-4	Sequence 4, Appli
51	32	61.5	529	4	US-09-240-639-4	Sequence 4, Appli
52	32	61.5	618	2	US-08-770-761A-3	Sequence 3, Appli
53	32	61.5	647	2	US-08-770-761A-8	Sequence 8, Appli
54	32	61.5	660	2	US-08-770-761A-2	Sequence 2, Appli
55	32	61.5	662	2	US-08-770-761A-5	Sequence 5, Appli
56	32	61.5	705	2	US-08-770-761A-7	Sequence 7, Appli
57	32	61.5	819	2	US-08-464-517-7	Sequence 7, Appli
58	32	61.5	819	3	US-08-246-361A-7	Sequence 7, Appli
59	32	61.5	819	3	US-08-463-772-7	Sequence 7, Appli
60	32	61.5	819	5	PCT-US93-05000-7	Sequence 7, Appli
61	31	59.6	59	4	US-08-963-851-14	Sequence 14, Appl
62	31	59.6	622	2	US-08-459-146-2	Sequence 2, Appli
63	31	59.6	622	2	US-08-459-065-2	Sequence 2, Appli
64	30	57.7	13	4	US-09-288-391-22	Sequence 22, Appl
65	30	57.7	13	4	US-09-288-391-23	Sequence 23, Appl
66	30	57.7	117	4	US-08-444-818-44	Sequence 44, Appl
67	30	57.7	121	4	US-09-152-060-68	Sequence 68, Appl
68	30	57.7	121	4	US-09-152-060-85	Sequence 85, Appl
69	30	57.7	122	4	US-08-879-995A-1	Sequence 1, Appli
70	30	57.7	122	3	US-09-215-096-1	Sequence 1, Appli
71	30	57.7	132	4	US-08-444-818-52	Sequence 52, Appl
72	30	57.7	159	2	US-08-844-086-4	Sequence 4, Appli
73	30	57.7	159	3	US-09-018-211-4	Sequence 4, Appli
74	30	57.7	211	5	PCT-US94-0417A-18	Sequence 18, Appl
75	30	57.7	241	3	US-08-834-776A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-228-986-73
; Sequence 73, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-73

Query Match 65.4%; Score 34; DB 4; Length 947;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMSYS 11
I:| | | | |
Db 686 VMPGSIYS 694

RESULT 2

US-08-637-759B-236
; Sequence 236, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-637-759B-236

Query Match 63.5%; Score 33; DB 2; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
| | | | |
Db 1 EEISPLGWSY 10

Query Match 63.5%; Score 33; DB 2; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
| | | | |
Db 1 EEISPLGWSY 10

RESULT 3

US-08-637-355A-236
; Sequence 236, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-236

Query Match 63.5%; Score 33; DB 3; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
| | | | |
Db 1 EEISPLGWSY 10

Query Match 63.5%; Score 33; DB 3; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
| | | | |
Db 1 EEISPLGWSY 10

RESULT 4

US-09-201-945-236
; Sequence 236, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/637,759

;
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMs 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-236

Query Match 63.5%; Score 33; DB 4; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
||: ||| |
Db 1 EEISPLGWSY 10

RESULT 5
517197-51
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 51:
; LENGTH: 65
5177197-51

Query Match 63.5%; Score 33; DB 6; Length 65;
Best Local Similarity 45.5%; Pred. No. 7.8;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
:|: | | | |
Db 52 KEICPGMGYT 62

RESULT 6
5177197-1
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 1:
; LENGTH: 410
5177197-1

Query Match 63.5%; Score 33; DB 6; Length 410;
Best Local Similarity 45.5%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
:|: | | | |
Db 399 KEICPGMGYT 409

RESULT 7
5177197-30
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 30:
; LENGTH: 1394
5177197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;
Best Local Similarity 45.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
:|: | | | |
Db 399 KEICPGMGYT 409

RESULT 8
US-09-357-952-66
; Sequence 66, Application US/09357952
; Patent No. 6248904
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Pro
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/09/357,952
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US 60/093,642
; EARLIER FILING DATE: 21-JUL-1998
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
:|: | | | |
Db 1 DDIVPCMSY 10

RESULT 9
US-09-521-650-66
; Sequence 66, Application US/09521650
; Patent No. 6335429
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard

; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/521,650
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 09/168,888
; EARLIER FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
Db 1 DDIVPCMSY 10

RESULT 10
US-09-168-888-66
; Sequence 66, Application US/09168888
; Patent No. 6342611
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/168,888
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
Db 1 DDIVPCMSY 10

RESULT 11
US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5856161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-1-Associated Protein Kinase And Methods
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5721CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHEICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
Db 24 EEVFLPMY 33

RESULT 12
US-08-460-694-4
; Sequence 4, Application US/08460694
; Patent No. 5858655
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,694
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McConathy, Evelyn H.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 0609.4070002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 20 EEVFFPLAMNY 29

RESULT 13
US-08-460-744-4
Sequence 4, Application US/08460744
Patent No. 6107541
GENERAL INFORMATION:
APPLICANT: Arnold, Andrew
TITLE OF INVENTION: Pradi Cyclin and its cdna
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,744
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McConathy, Evelyn H.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 0609.4070005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-744-4

Query Match 61.5%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 20 EEVFFPLAMNY 29

RESULT 14
US-07-667-711B-4
Sequence 4, Application US/07667711B
Patent No. 6110700
GENERAL INFORMATION:
APPLICANT: ARNOLD, ANDREW
TITLE OF INVENTION: Pradi Cyclin and its cdna
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/667,711B
FILING DATE: 11-MAR-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MCPHAIL, DONALD R.
REGISTRATION NUMBER: 35,811
REFERENCE/DOCKET NUMBER: 0609.4070000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 20 EEVFFPLAMNY 29

RESULT 15
US-08-193-977-7
Sequence 7, Application US/08193977
Patent No. 5625031
GENERAL INFORMATION:
APPLICANT: WEBSTER, KEVIN R.
APPLICANT: COLEMAN, KEVIN G.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND

;; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN
;; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN
;; NUMBER OF SEQUENCES: 34
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: REED & ROBINS
;; STREET: 635 BRYANT STREET
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: UNITED STATES OF AMERICA
;; ZIP: 94301

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/193,977
;; FILING DATE: 08-FEB-1994
;; CLASSIFICATION: 530

;; ATTORNEY/AGENT INFORMATION:
;; NAME: ROBINS, ROBERTA L.
;; REGISTRATION NUMBER: 33,208
;; REFERENCE/DOCKET NUMBER: 5998-0016
;; TELEPHONE: (415) 617-8999
;; TELEFAX: (415) 327-3231

;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 173 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide

US-08-193-977-7

Query Match 61.5%; Score 32; DB 1; Length 173;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
DB 55 EEVFPPLAMNY 64

RESULT 16
US-08-464-517-21
;; Sequence 21, Application US/08464517
;; Patent No. 5869640
;; GENERAL INFORMATION:
;; APPLICANT: BEACH, David H.
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII(text)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/464,517
;; FILING DATE:
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/963,308
;; FILING DATE: 16-OCT-1992
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-464-517-21

Query Match 61.5%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
DB 74 EEVFPPLAMNY 83

RESULT 17
US-08-246-361A-21
;; Sequence 21, Application US/08246361A
;; Patent No. 5998582
;; GENERAL INFORMATION:
;; APPLICANT: BEACH, David H.
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII(text)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/246,361A
;; FILING DATE: 19-MAY-1994
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/963,308
;; FILING DATE: 16-OCT-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941

;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-21

Query Match 61.5%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFPPLAMNY 83

RESULT 18

US-08-463-772-21
; Sequence 21, Application US/08463772
; Patent No 6066501

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent

; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 189 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-463-772-21

Query Match 61.5%; Score 32; DB 3; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFPPLAMNY 83

RESULT 19

PCT-US93-05000-21
; Sequence 21, Application PC/TUS9305000

; GENERAL INFORMATION:

; APPLICANT: MITOTIX

; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: US

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/05000

; FILING DATE: 19930525

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/888,178

; FILING DATE: 26-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: CSHL91-02A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 616-861-9540

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 189 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

PCT-US93-05000-21

Query Match

61.5%; Score 32; DB 5; Length 189;

Best Local Similarity 60.0%; Pred. No. 42;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

||| | | |

Db 74 EEVFPPLAMNY 83

RESULT 20

US-08-464-517-22

; Sequence 22, Application US/08464517

; Patent No. 5869640

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/464,517

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-22

Query Match 61.5%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 20 EEVFPPLAMNY 29

RESULT 21

US-08-246-361A-22
; Sequence 22, Application US/08246361A
; Patent No. 5998582

; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA

; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-22

Query Match 61.5%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 20 EEVFPPLAMNY 29

RESULT 22

US-08-463-772-22
; Sequence 22, Application US/08463772
; Patent No. 6066501

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,772

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/888,178

; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514

; FILING DATE: 16-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Matthew P. Vincent

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 236 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-463-772-22

Query Match 61.5%; Score 32; DB 3; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 20 EEVFPPLAMNY 29

RESULT 23
PCT-US93-05000-22
; Sequence 22, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-05000-22
Query Match 61.5%; Score 32; DB 5; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
Db 20 EEVFLPMNY 29

RESULT 24
US-08-464-517-6
; Sequence 6, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIL-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-517-6
Query Match 61.5%; Score 32; DB 2; Length 280;
Best Local Similarity 60.0%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
Db 75 EEVFLPMNY 84

RESULT 25
US-08-463-772-6
; Sequence 6, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIL-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-772-6

Query Match 61.5%; Score 32; DB 3; Length 280;
Best Local Similarity 60.0%; Pred. NO. 65;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | |:
Db 75 EEVFPPLAMNY 84

Search completed: June 10, 2003, 13:51:38
Job time : 9.64286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 seconds
(without alignments)
75.710 Million cell updates/sec

Title: US-09-909-164-47
Perfect score: 52
Sequence: 1 BEVVPXGMSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues
Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US03_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	3472	9	US-10-027-806-4
2	38	73.1	3472	9	US-10-034-623-4
3	38	73.1	3472	9	US-10-027-801-4
4	34	65.4	947	9	US-10-101-464A-73
5	33	63.5	426	9	US-10-214-766-43
6	33	63.5	478	9	US-09-924-340-108
7	33	63.5	478	9	US-09-992-600A-108
8	33	63.5	478	9	US-09-746-783-184
9	33	63.5	478	9	US-10-000-489-108
10	33	63.5	478	9	US-10-000-986-108
11	33	63.5	653	9	US-09-820-843A-26
12	32	61.5	254	10	US-09-947-387-66
13	32	61.5	289	9	US-09-778-927A-53
14	32	61.5	289	9	US-10-024-066-2
15	32	61.5	289	9	US-10-024-066-4
16	32	61.5	289	10	US-09-919-497-54
17	32	61.5	295	10	US-09-925-300-1061
18	32	61.5	529	10	US-09-923-304-4
19	32	61.5	691	9	US-10-101-921-4

ALIGNMENTS

RESULT 1

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAUM SYMBIOSUM
; FILE REFERENCE: DCCORP.002A
; CURRENT APPLICATION NUMBER: US/10/027,806
; CURRENT FILING DATE: 2001-12-21

20	61.5	691	10	US-09-925-731-2	Sequence 2, Appl1
21	61.5	1377	10	US-09-815-242-10384	Sequence 10384, A
22	61.5	2799	9	US-10-151-736-4	Sequence 4, Appl1
23	59.6	53	9	US-10-092-154-878	Sequence 878, App
24	59.6	53	10	US-09-764-847-878	Sequence 14, Appl
25	59.6	59	10	US-09-948-080-14	Sequence 14, Appl
26	59.6	161	9	US-09-738-626-5124	Sequence 5124, Ap
27	59.6	163	9	US-10-117-846-20	Sequence 20, Appl
28	59.6	192	9	US-09-986-480-171	Sequence 171, App
29	59.6	198	10	US-09-731-872-334	Sequence 334, App
30	59.6	223	9	US-09-738-626-6349	Sequence 6349, Ap
31	59.6	299	10	US-09-815-242-10697	Sequence 10697, A
32	59.6	381	9	US-09-975-139-5	Sequence 5, Appl1
33	59.6	702	9	US-10-280-403-2	Sequence 2, Appl1
34	59.6	702	10	US-09-907-479-2	Sequence 2, Appl1
35	59.6	763	9	US-09-738-626-4454	Sequence 4454, Ap
36	59.6	1053	10	US-09-815-242-5136	Sequence 5136, Ap
37	59.6	1407	10	US-09-815-242-10439	Sequence 10439, A
38	59.6	1426	10	US-09-912-020-340	Sequence 340, App
39	57.7	7	9	US-09-909-062-1	Sequence 1, Appl1
40	57.7	7	9	US-09-909-062-9	Sequence 9, Appl1
41	57.7	7	9	US-09-909-062-130	Sequence 130, App
42	57.7	121	9	US-09-852-797-68	Sequence 68, Appl
43	57.7	121	9	US-09-852-797-85	Sequence 85, Appl
44	57.7	121	10	US-09-853-161-68	Sequence 68, Appl
45	57.7	121	10	US-09-853-161-85	Sequence 85, Appl
46	57.7	121	10	US-09-852-659A-68	Sequence 68, Appl
47	57.7	121	10	US-09-852-659A-85	Sequence 85, Appl
48	57.7	135	9	US-09-992-598-359	Sequence 359, App
49	57.7	135	9	US-09-989-293A-359	Sequence 359, App
50	57.7	135	9	US-09-989-735-359	Sequence 359, App
51	57.7	135	9	US-09-990-444-359	Sequence 359, App
52	57.7	135	9	US-09-989-730-359	Sequence 359, App
53	57.7	135	9	US-09-990-436-359	Sequence 359, App
54	57.7	135	9	US-09-991-181-359	Sequence 359, App
55	57.7	135	9	US-09-993-687-359	Sequence 359, App
56	57.7	135	9	US-09-989-734-359	Sequence 359, App
57	57.7	135	9	US-09-997-653-359	Sequence 359, App
58	57.7	135	9	US-10-174-590-444	Sequence 444, App
59	57.7	135	9	US-10-176-758-444	Sequence 444, App
60	57.7	135	9	US-10-175-737-444	Sequence 444, App
61	57.7	135	9	US-09-993-667-359	Sequence 359, App
62	57.7	135	9	US-10-173-706-444	Sequence 444, App
63	57.7	135	9	US-10-175-738-444	Sequence 444, App
64	57.7	135	9	US-10-175-752-444	Sequence 444, App
65	57.7	135	9	US-10-176-482-444	Sequence 444, App
66	57.7	135	9	US-10-176-757-444	Sequence 444, App
67	57.7	135	9	US-10-176-913-444	Sequence 444, App
68	57.7	135	9	US-10-180-552-444	Sequence 444, App
69	57.7	135	9	US-10-180-557-444	Sequence 444, App
70	57.7	135	9	US-09-990-438-359	Sequence 359, App
71	57.7	135	9	US-09-990-562-359	Sequence 359, App
72	57.7	135	9	US-09-997-428-359	Sequence 359, App
73	57.7	135	9	US-09-997-666-359	Sequence 359, App
74	57.7	135	9	US-10-173-700-444	Sequence 444, App
75	57.7	135	9	US-10-174-572-444	Sequence 444, App

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
Db 2294 EDVIPRGISFS 2304
|:|:| |:|:|

RESULT 2
US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
Db 2294 EDVIPRGISFS 2304
|:|:| |:|:|

RESULT 3
US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/027,801
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match 73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11
Db 2294 EDVIPRGISFS 2304
|:|:| |:|:|

RESULT 4
US-10-101-464A-73
; Sequence 73, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-73

Query Match 65.4%; Score 34; DB 9; Length 947;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
Db 686 VMPGSGISYS 694
|:| |:|:|

RESULT 5
US-10-214-766-43
; Sequence 43, Application US/10214766
; Publication No. US20030084473A1
; GENERAL INFORMATION:
; APPLICANT: Gocal, Greg
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
; FILE REFERENCE: CA1138
; CURRENT APPLICATION NUMBER: US/10/214,766
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,734
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 43
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-214-766-43

Query Match 63.5%; Score 33; DB 9; Length 426;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10

Db 223 EFVIPAGQSY 232
| | | | |

RESULT 6
US-09-924-340-108
; Sequence 108, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-108

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
| | | | |
Db 239 EVAPAGASYN 248

RESULT 7
US-09-992-600A-108
; Sequence 108, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 EVVPXGMSYS 11
| | | | |
Db 239 EVAPAGASYN 248

RESULT 8
US-09-746-783-184
; Sequence 184, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 231
; ENCODING THEM
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746,783
; FILING DATE: 21-Dec-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Milasincic, Debra J.
; REGISTRATION NUMBER: 46,931
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 184:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 184:
US-09-746-783-184

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
| | | | |
Db 239 EVAPAGASYN 248

RESULT 9
US-10-000-489-108
; Sequence 108, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

RESULT 11

```
QY      1 EEVVPXGMSY 10
        ::::| | |
Db      1 DDIVPCSMYS 10
```

RESULT 13
US-09-778-927A-53
; Sequence 53, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(254)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-53

Query Match 61.5%; Score 32; DB 10; Length 254;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFFPLAMNY 83

RESULT 14
US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-024-066-2

Query Match 61.5%; Score 32; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 73 EEVFFPLAMNY 82

RESULT 15
US-10-024-066-4

; Sequence 4, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-066-4

Query Match 61.5%; Score 32; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFFPLAMNY 83

RESULT 16
US-09-919-497-54
; Sequence 54, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-54

Query Match 61.5%; Score 32; DB 10; Length 289;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFFPLAMNY 83

RESULT 17
US-09-925-300-1061
; Sequence 1061, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (243)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (277)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1061

Query Match 61.5%; Score 32; DB 10; Length 295;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
||| |||
Db 52 EVLPKMSYA 61

RESULT 18
US-09-923-304-4
; Sequence 4, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:658US
; CURRENT APPLICATION NUMBER: US/09/923,304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-304-4

Query Match 61.5%; Score 32; DB 10; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
:||:| |:
Db 48 QEVLPPGLKY 57

RESULT 19
US-10-101-921-4
; Sequence 4, Application US/10101921
; Publication No. US20030022199A1
; GENERAL INFORMATION:
; APPLICANT: Nezu, Jun-Ichi
; APPLICANT: Ose, Asuka
; APPLICANT: Tsuji, Akira
; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E
; FILE REFERENCE: 06501-104US1
; CURRENT APPLICATION NUMBER: US/10/101,921
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: PCT/JP00/06416
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: JP 11/267835
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-921-4

Query Match 61.5%; Score 32; DB 9; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 10
:|| |:
Db 188 IVPLGLSY 195

RESULT 20
US-09-925-731-2
; Sequence 2, Application US/09925731
; Patent No. US20020090622A1
; GENERAL INFORMATION:
; APPLICANT: ADEKUN, ANTHONI MONISOLA
; APPLICANT: AMBROSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DJB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925,731
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match 61.5%; Score 32; DB 10; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 10
:|| |:
Db 188 IVPLGLSY 195

RESULT 21
US-09-815-242-10384
; Sequence 10384, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10384
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10384

Query Match 61.5%; Score 32; DB 10; Length 1377;
Best Local Similarity 45.5%; Pred. No. 6.7e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSYS 11
||| :|||
Db 369 EQLNPAGLSYT 379

RESULT 22
US-10-151-736-4
; Sequence 4, Application US/10151736
; Publication No. US20020192160A1
; GENERAL INFORMATION:
; APPLICANT: Callaghan, Michelle J.
; APPLICANT: Sutherland, Linfield
; APPLICANT: Watts, Colin K.
; TITLE OF INVENTION: No. US20020192160A1 Human Tumour Suppressor Gene
; FILE REFERENCE: RICE-010CON
; CURRENT APPLICATION NUMBER: US/10/151,736
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/403,402
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: PCT/AU98/00280
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2799
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-736-4

Query Match 61.5%; Score 32; DB 9; Length 2799;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
||| :|||
Db 2096 EVLPTRMSYA 2105

RESULT 23
US-10-092-154-878
; Sequence 878, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-878

Query Match 59.6%; Score 31; DB 9; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| :|||
Db 39 VVPTAVSYS 47

RESULT 24
US-09-764-847-878
; Sequence 878, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-878

Query Match 59.6%; Score 31; DB 10; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| :|||
Db 39 VVPTAVSYS 47

RESULT 25
US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102702A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match 59.6%; Score 31; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 BEVVPXGMSYS 11
I: I I: I
Db 38 EKHIPGGLEYS 48

Search completed: June 10, 2003, 14:35:46
Job time : 15.0714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds
(without alignments)
94.297 Million cell updates/sec

Title: US-09-909-164-47
Perfect score: 52
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	73.1	3472	2 T31308	hypothetical 367K
2	37	71.2	840	2 T39116	probable sulfate p
3	37	71.2	877	2 T40413	sulfate permease -
4	36	69.2	102	2 A42452	V1 protein - tobac
5	36	69.2	1498	2 B97355	DNA segregation AT
6	35	67.3	225	2 S57810	hypothetical prote
7	35	67.3	425	2 T24111	hypothetical prote
8	35	67.3	670	2 S22293	zinc finger protei
9	35	67.3	749	2 H82691	topoisomerase IV s
10	35	67.3	2717	2 A34203	DNA-binding protei
11	34	65.4	156	2 S54619	hypothetical prote
12	34	65.4	252	2 H69491	cell division inhi
13	34	65.4	544	2 C82900	probable ABC subst
14	33	63.5	94	2 T40758	hypothetical prote
15	33	63.5	116	2 E90544	50S ribosomal prot
16	33	63.5	165	2 D69493	hypothetical prote
17	33	63.5	253	2 C81374	hypothetical prote
18	33	63.5	259	2 T34536	hypothetical prote
19	33	63.5	284	2 S75817	hypothetical prote
20	33	63.5	298	2 T47670	beta-ketoacyl-ACP
21	33	63.5	368	2 F72281	hypothetical prote
22	33	63.5	426	2 D82163	3-phosphoshikimate
23	33	63.5	466	2 T43653	cdc37 protein - fi
24	33	63.5	653	2 D82352	iron(III) ABC tran
25	33	63.5	890	2 A30481	bactriocin BCN5 -
26	33	63.5	1028	2 A34286	ATP-dependent DNA
27	33	63.5	1152	2 D87046	conserved hypothet
28	33	63.5	1394	2 A35626	transforming growt
29	33	63.5	1401	2 G82336	DNA-directed RNA p

ALIGNMENTS

RESULT 1

T31308
hypothetical 367K protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000

R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

A:Title: Genomic analysis reveals chromosomal variation in natural populations of

A:Reference number: Z20994; MUID:98422450; PMID:9748430

A:Accession: T31308

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3472 <SCH>

A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AAC62699.1

C:Supfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 73.1%; Score 38; DB 2; Length 3472;
Best Local Similarity 54.5%; Pred. No. 59;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

hypothetical prote
masking protein pr
hypothetical prote
transport protein
rho protein GPP-di
hypothetical prote
cyclin D2 - rat
cyclin D2 - mouse
cyclin D2 - human
cyclin D1 - Africa
cyclin D2 - Africa
cyclin D2 - chicke
cyclin D1 - zebra
cyclin D3 - human
cyclin D1 - human
cyclin D1 - mouse
cyclin D1 - rat
rhesF protein - Esc
conserved hypothet
tolB protein - Hae
conserved hypothet
conserved hypothet
agaa protein limpo
hypothetical prote
transport protein
infected cell prot
SCT1 protein - yea
hypothetical prote
maltooligosyltreha
RNA 1 protein - to
aggregation protei
rhaA protein precu
rhaA protein in rh
RhsH core protein
RhsC protein in rh
RhsC protein precu
RhsC core protein
RhsB protein precu
gene 11-1 protein
trans-regulatory s
hypothetical prote
probable cobH - My
probable purQ prot
phosphoribosylform

33 63.5 1548 2 T04456
33 63.5 1712 2 A38261
32 61.5 84 2 E97333
32 61.5 175 2 P00616
32 61.5 223 2 T01457
32 61.5 279 2 B72481
32 61.5 288 2 JC4011
32 61.5 288 2 I58372
32 61.5 289 2 A41984
32 61.5 289 2 A42822
32 61.5 291 2 S57922
32 61.5 291 2 S57925
32 61.5 291 2 JC4579
32 61.5 291 2 S62730
32 61.5 292 2 B42822
32 61.5 295 2 A38977
32 61.5 295 2 A56523
32 61.5 295 2 JC2342
32 61.5 347 2 I55120
32 61.5 363 2 D69551
32 61.5 427 2 F64064
32 61.5 427 2 A9286
32 61.5 449 2 B0604
32 61.5 498 2 B0604
32 61.5 525 2 D98311
32 61.5 525 2 AF2971
32 61.5 726 2 T44000
32 61.5 726 2 T44187
32 61.5 759 2 S25330
32 61.5 889 2 S22659
32 61.5 922 2 AG1827
32 61.5 993 1 P1VXTA
32 61.5 1306 2 S22624
32 61.5 1377 2 C65159
32 61.5 1377 2 E86034
32 61.5 1394 2 H91236
32 61.5 1397 2 A85570
32 61.5 1397 2 C64805
32 61.5 1399 2 A99720
32 61.5 1409 2 F91187
32 61.5 1411 2 E65145
32 61.5 1948 2 S00485
31 59.6 124 1 VKLJ51
31 59.6 133 2 A71173
31 59.6 208 1 D70764
31 59.6 224 2 G70709
31 59.6 224 2 F87186

Db 2294 EDVIPRGISFS 2304

Query Match 71.2%; Score 37; DB 2; Length 840;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| |||||
Db 135 VVPOGMSYA 143

RESULT 3
T40413
sulfate permease - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40413
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21926
A:Accession: T40413
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-840 <HUN>
A:Cross-references: EMBL:AL132779; PTDN:CAB60015.1; GSPDB:GN000066; SPDB:SPAC869.05C
A:Experimental source: strain 972h-; cosmid c869
C:Genetics:
A:Gene: SPDB:SPAC869.05C
A:Map position: 1

Query Match 71.2%; Score 37; DB 2; Length 877;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| |||||
Db 148 VVPOGMSYA 156

RESULT 4
A42452
V1 protein - tobacco yellow dwarf virus (strain Australia)
C:Species: tobacco yellow dwarf virus
C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
C:Accession: A42452
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
Virology 187, 633-642, 1992
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow
A:Reference number: A42452; MUID:92188538; PMID:1546458
A:Accession: A42452
A:Molecule type: DNA
A:Residues: 1-102 <MOR>

Db 2294 EDVIPRGISFS 2304

Query Match 69.2%; Score 36; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
||| |||||
Db 7 QVVPXGMSYS 16

RESULT 5
B97355
DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported]
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97355
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97355
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1498 <KUR>
A:Cross-references: GB:AE001437; PTDN:AAK81629.1; PID:g15026814; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3709

Query Match 69.2%; Score 36; DB 2; Length 1498;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 10
|:|:| |||||
Db 1276 EQKIPGMSY 1285

RESULT 6
S57810
hypothetical protein precursor (clone TP11) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S57810
R:Milligan, S.B.; Gasser, C.S.
Plant Mol. Biol. 28, 691-711, 1995
A:Title: Nature and regulation of pistil-expressed genes in tomato.
A:Reference number: S57808; MUID:95375233; PMID:7647301
A:Accession: S57810
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-225 <ML>
A:Cross-references: EMBL:U20592; NID:g924625; PTDN:AAA80497.1; PID:g924626
C:Superfamily: plant kunitz-type proteinase inhibitor

Query Match 67.3%; Score 35; DB 2; Length 225;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
:|||| |:
Db 32 DEWVPNGKTYA 42

RESULT 7
T24111
hypothetical protein R10D12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24111
R:Percy, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: 219842
 A:Accession: T24111
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-425 <WIL>
 A:Cross-references: EMBL:281109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10
 A:Experimental source: Clone R10D12
 C:Genetics:
 A:Gene: CESP:R10D12.10
 A:Map position: 5
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 67.3%; Score 35; DB 2; Length 425;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 I::I I: I
 Db 335 EQIVPGGLQY 344

RESULT 8

S22293
 zinc finger protein At-BP2 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Sep-1999
 C:Accession: S22293; I78656
 R:Mitchellmore, C.; Traboni, C.; Cortese, R.
 Nucleic Acids Res. 19, 141-147, 1991
 A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-Nucleic Acids Res. 19, 141-147, 1991
 A:Reference number: I58280; MUID:91187610; PMID:1901405
 A:Accession: S22293

A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-670 <MIT>
 A:Cross-references: EMBL:X54250; NID:g57519; PIDN:CAA38151.1; PID:g57520
 A:Note: the authors did not translate the codon for residue 1
 C:Superfamily: HIV-EP2 enhancer-binding protein
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 I::I I: I
 Db 376 VVPAGLTYS 384

RESULT 9

H82691
 topoisomerase IV subunit XF153 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: H82691

R:anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: GB:A2515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82691

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-749 <SIM>

A:Cross-references: GB:AE003967; GB:AE003849; NID:g9106347; PIDN:AAF84162.1; GSPDB:GN001

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carrier, H

sub-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Mart
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki,
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmier
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; S
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da S
 M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF153

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisome

Query Match 67.3%; Score 35; DB 2; Length 749;

Best Local Similarity 77.8%; Pred. No. 48;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
 I::I I: I
 Db 526 EVDPGMSY 534

RESULT 10

A34203

DNA-binding protein PRDII-BF1 - human

N:Alternate names: major histocompatibility complex enhancer-binding protein 1

C:Species: Homo sapiens (man)

C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 20-Sep-1999

C:Accession: A34203; A34779

R:Fan, C.M.; Maniatis, T.

Genes Dev. 4, 29-42, 1990

A:Title: A DNA-binding protein containing two widely separated zinc finger motifs

A:Reference number: A34203; MUID:90169514; PMID:2106471

A:Accession: A34203

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2717 <FAN>

A:Cross-references: EMBL:X51435; NID:g38017; PIDN:CAA35798.1; PID:g38018

R:Baldwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.

Mol. Cell. Biol. 10, 1406-1414, 1990

A:Title: A large protein containing zinc finger domains binds to related sequence

A:Reference number: A34779; MUID:90205817; PMID:2108316

A:Accession: A34779

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 801-1072, 'N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 1

A:Cross-references: GB:M32019

C:Superfamily: HIV-EP2 enhancer-binding protein

C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 I::I I: I
 Db 2405 VVPAGLTYS 2413

RESULT 11

S54619

hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C:Accession: S54619; S66879

R:de Haan, M.; Maarse, A.C.; Grivell, L.A.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54619

A:Accession: S54619

A:Molecule type: DNA

A:Residues: 1-156 <DEH>

A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829133

R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66877

A:Accession: S66879

A:Molecule type: DNA

A:Residues: 1-156 <DEW>

A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR013

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0005539

A:Map position: 15R

C:Superfamily: hypothetical protein YOR013w

Query Match 65.4%; Score 34; DB 2; Length 156;
 Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10

Db 50 EVPLGMDY 58

RESULT 12

H69491

cell division inhibitor (mind-2) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 26-Aug-1999

C:Accession: H69491

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69491

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-252 <KLE>

A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AA889318.1; PID:g264860

C:Superfamily: cell division inhibitor mind

Query Match 65.4%; Score 34; DB 2; Length 252;

Best Local Similarity 75.0%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9

Db 81 EVIPAGMS 88

RESULT 13

C82900

probable ABC substrate-binding protein, iron UU359 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: C82900

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: C82900

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <GLA>

A:Cross-references: GB:AF002133; GB:AF222894; NID:g6899339; PIDN:AAF30768.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: ABCsbp-5; UU359

A:Genetic code: SGC3

Query Match 65.4%; Score 34; DB 2; Length 544;

Best Local Similarity 70.0%; Pred. No. 55;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db 135 EEVVPYLSY 144

RESULT 14

I40758

hypothetical protein 1 - Campylobacter jejuni (fragment)

C:Species: Campylobacter jejuni

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 08-Oct-1999

C:Accession: I40758; S47317

R:Hani, E.K.; Chan, V.L.

J. Bacteriol. 177, 2396-2402, 1995

A:Title: Expression and characterization of Campylobacter jejuni benzoylglycine am

A:Reference number: I40758

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-94 <RES>

A:Cross-references: EMBL:Z36940; NID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match 63.5%; Score 33; DB 2; Length 94;

Best Local Similarity 55.6%; Pred. No. 13;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10

Db 26 DIPFPGMSY 34

RESULT 15

E90544

SOS ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Mosze

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma

A:Reference number: A99512; MUID:21267165; PMID:11333084

A:Accession: E90544

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <KUR>

A:Cross-references: GB:AL445566; PID:g14089674; PIDN:CAC13434.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPU_2610

A:Genetic code: SGC3

C:Superfamily: Escherichia coli ribosomal protein L20

Query Match 53.5%; Score 33; DB 2; Length 116;

Best Local Similarity 77.8%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db 68 VRPLGMSYS 76

RESULT 16

D69493

hypothetical protein AF1949 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 04-Mar-2000

C:Accession: D69493

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; I

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Syk

Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: D69493

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-165 <KLE>

A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AAB89307.1; PID:g264859

C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1949

Query Match 63.5%; Score 33; DB 2; Length 165;

Best Local Similarity 60.0%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10

||| ||| ||

Db 60 EESIPDGASY 69

RESULT 17

C81374

hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: C81374

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: C81374

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73246.1; PID:g696842

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;

Best Local Similarity 55.6%; Pred. No. 39;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGMSY 10

::| ||||

Db 185 DIFPSGMSY 193

RESULT 18

T34536

hypothetical protein DKFZp434C031.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34536

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34536

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-259 <POU>

A:Cross-references: EMBL:AL122063

A:Experimental source: adult testis; clone DKFZp434C031

C:Genetics:

A:Note: DKFZp434C031.1

Query Match 63.5%; Score 33; DB 2; Length 259;

Best Local Similarity 60.0%; Pred. No. 40;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EVVPXGMSY 11

||| ||| ||

Db 22 EVAPAGASYN 31

RESULT 19

S75817

hypothetical protein slr1275 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S75817

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechoc

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75817

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-284 <KAN>

A:Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAA18276.1; PID:d

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.5%; Score 33; DB 2; Length 284;

Best Local Similarity 55.6%; Pred. No. 44;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11

||| ||||

Db 208 VIPAGVSYT 216

RESULT 20

T47670

beta-ketoacyl-ACP reductase-like protein - Arabidopsis thaliana

N:Alternate names: protein T26112.190

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 18-Aug-2000

C:Accession: T47670

R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer,

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24471

A:Accession: T47670

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <MON>

A:Cross-references: EMBL:AL132954

A:Experimental source: cultivar Columbia; BAC clone T26112

C:Genetics:

A:Map position: 3

A:Introns: 25/3

A:Note: T26112.190

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 63.5%; Score 33; DB 2; Length 298;

Best Local Similarity 55.6%; Pred. No. 47;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11

||| ||||

Db 187 IVPGGLAYS 195

RESULT 21

F72281

hypothetical protein TM1216 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: F72281

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from geno

A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: F72281
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <ARN>
A:Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36291.1; PID:g498176
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1216
C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

Query Match 63.5%; Score 33; DB 2; Length 368;
Best Local Similarity 55.6%; Pred. No. 59;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
:|||
Db 294 IVPKGMAYA 302

RESULT 22

D82163
3-phosphoshikimate 1-carboxyvinyltransferase VC1732 [imported] - Vibrio cholerae (strain D82163)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82163
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82163
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <HE>
A:Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94882.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1732
A:Map position: 1
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

Query Match 63.5%; Score 33; DB 2; Length 426;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
:|||
Db 223 EFVIPAGQSY 232

RESULT 23

T43653
cdc37 protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T43653; T40791; T43654
R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.
submitted to the EMBL Data Library, March 1999
A:Description: Schizosaccharomyces pombe cdc37 cDNA.
A:Reference number: 222602
A:Accession: T43653
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-466 <WES>
A:Cross-references: EMBL:AJ132376; PIDN:CAB38757.1
R:Wood, V.; Raandream, M.A.; Barrall, B.G.; Volckaert, G.
submitted to the EMBL Data Library, May 1999
A:Reference number: 221875
A:Accession: T40791
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-466 <WOO>
A:Cross-references: EMBL:AL049769; PIDN:CAB42371.2; GSPDB:GN00067
A:Experimental source: strain 972h-; cosmid c9B6
R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.
submitted to the EMBL Data Library, March 1999
A:Description: Schizosaccharomyces pombe cdc37 gene.
A:Reference number: 222603
A:Accession: T43654
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-466 <WE2>
A:Cross-references: EMBL:AJ132377; PIDN:CAB38758.1
C:Genetics:
A:Gene: cdc37; SPAC9B6.10
A:Map position: 2
A:Introns: 8/2; 17/2; 21/1

Query Match 63.5%; Score 33; DB 2; Length 466;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
:|||
Db 98 DSAIPGMSY 107

RESULT 24

D82352
Iron(III) ABC transporter, permease protein VC0203 [imported] - Vibrio cholerae (strain D82352)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82352
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82352
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-653 <HEI>
A:Cross-references: GB:AE004110; GB:AE003852; NID:g9654600; PIDN:AAF93379.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0203
A:Map position: 1

Query Match 63.5%; Score 33; DB 2; Length 653;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9
:|||||
Db 300 EEVVPSGIT 308

RESULT 25

A30481
bacteriocin BCN5 - Clostridium perfringens plasmid pIP404
C:Species: Clostridium perfringens
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Nov-1999
C:Accession: A30481; S03779
R:Garnier, T.; Cole, S.T.
J. Bacteriol. 168, 1189-1196, 1986
A:Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens
A:Reference number: J70354; MUID:87057020; PMID:2877971
A:Accession: A30481
A:Molecule type: DNA
A:Residues: 1-890 <GAR>
A:Cross-references: GB:M32882; GB:J03309; NID:g150738; PIDN:AAA98249.1; PID:g150738
C:Genetics:
A:Gene: bcn

A:Genome: plasmid
C:Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5
C:Keywords: bacteriocin

Query Match 63.5%; Score 33; DB 2; Length 890;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
| | | | | : |
Db 170 EVVPGGFTY 178

Search completed: June 10, 2003, 13:49:19
Job time : 12.2143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds
(without alignments)
101.387 Million cell updates/sec

Title: US-09-909-164-47
Perfect score: 52
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8986 fusobacteri
2	37	71.2	877	1 SULH_SCHPO	O74377 schizosacch
3	36	69.2	102	1 YILK_TYDVA	P31619 tobacco yel
4	36	69.2	1498	1 YIA9_CLOAB	Q04351 clostridium
5	35	67.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
6	34	65.4	788	1 CY14_NEUCR	P23622 neurospora
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q98qvo mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIECH	Q9kro0 vibrio chol
11	33	63.5	466	1 CC37_SCHPO	O94740 schizosacch
12	33	63.5	478	1 GSR2_HUMAN	Q9nzm5 homo sapien
13	33	63.5	890	1 BCN5_CLOPE	P08696 clostridium
14	33	63.5	1394	1 LTBS_HUMAN	P22064 homo sapien
15	33	63.5	1401	1 RPOC_VIECH	Q9kv29 vibrio chol
16	33	63.5	1595	1 LTBI_HUMAN	Q14766 homo sapien
17	33	63.5	1712	1 LTBI_RAT	Q00918 rattus norv
18	32	61.5	288	1 CGD2_RAT	Q04827 rattus norv
19	32	61.5	289	1 CGD2_HUMAN	Q30279 homo sapien
20	32	61.5	289	1 CGD2_MOUSE	P30280 mus musculus
21	32	61.5	291	1 CGD1_BRARE	Q90459 brachydanio
22	32	61.5	291	1 CGD1_XENLA	P50755 xenopus lae
23	32	61.5	291	1 CGD2_CHICK	P49706 gallus gall
24	32	61.5	291	1 CGD2_XENLA	P53782 xenopus lae
25	32	61.5	292	1 CGD1_CHICK	P55169 gallus gall
26	32	61.5	292	1 CGD3_HUMAN	P30281 homo sapien
27	32	61.5	295	1 CGD1_HUMAN	P24385 homo sapien
28	32	61.5	295	1 CGD1_MOUSE	P25322 mus musculus
29	32	61.5	295	1 CGD1_RAT	P39948 rattus norv
30	32	61.5	427	1 TOLB_HAEIN	P44677 haemophilus
31	32	61.5	529	1 ENP3_HUMAN	O75355 homo sapien
32	32	61.5	691	1 OAT6_HUMAN	Q9Y616 homo sapien
33	32	61.5	726	1 PRTP_HSV6U	P52384 human herpe

34	61.5	759	1	SCT1_YEAST	P32784 saccharomyc
35	61.5	920	1	EDD_RAT	Q62671 rattus norv
36	61.5	993	1	VIA_TAV	P28931 tomato aspe
37	61.5	1377	1	RHSA_ECOLI	P16916 escherichia
38	61.5	1397	1	RHSC_ECOLI	P16918 escherichia
39	61.5	1411	1	RHSE_ECOLI	P16917 escherichia
40	61.5	2799	1	EDD_HUMAN	O95071 homo sapien
41	59.6	124	1	REV_SIVCZ	P17280 chimpanzee
42	59.6	208	1	COBH_MYCTU	Q10676 mycobacteri
43	59.6	223	1	PURQ_PYRHO	O59619 pyrococcus
44	59.6	224	1	PURQ_HALN1	Q9hnu2 halobacteri
45	59.6	224	1	PURQ_MYCLE	O05756 mycobacteri
46	59.6	224	1	PURQ_MYCTU	P1841 mycobacteri
47	59.6	225	1	PURQ_CORAM	Q9rhxo corynebacte
48	59.6	240	1	GDIR_ARATH	Q9sf6c arabidopsi
49	59.6	276	1	Y939_METJA	Q58349 methanococc
50	59.6	319	1	YHAI_CRYPA	P10941 cryptoneutr
51	59.6	432	1	PURA_YEAST	P80210 saccharomyc
52	59.6	488	1	NOM2_PSEAE	Q9htr0 pseudomonas
53	59.6	670	1	OATP_RAT	P46720 rattus norv
54	59.6	706	1	ADDG_HUMAN	Q9uey8 homo sapien
55	59.6	706	1	ADDG_MOUSE	Q9qyb5 mus musculu
56	59.6	827	1	PLSB_MOUSE	Q61586 mus musculu
57	59.6	828	1	PLSB_RAT	P97564 rattus norv
58	59.6	1047	1	EF3_SCHPO	O94489 schizosacch
59	59.6	1407	1	RPOC_ECOLI	P00577 escherichia
60	59.6	1426	1	RHSD_ECOLI	P16919 escherichia
61	59.6	2145	1	U520_CAEEL	Q9u290 caenorhabdi
62	58.7	472	1	ET2A_XENLA	P19102 xenopus lae
63	57.7	81	1	YE47_ARCFU	O28825 archaeoglob
64	57.7	121	1	TKNK_HUMAN	Q8uhf0 homo sapien
65	57.7	132	1	ATPE_AQUAE	O86903 aquifex ae
66	57.7	146	1	ATPE_LACAC	Q99y0 lactobacill
67	57.7	218	1	PURQ_METH	O36270 methanobact
68	57.7	223	1	PURQ_RHIME	Q92p11 rhizobium m
69	57.7	223	1	RPIA_BUCAI	P57489 buchnera ap
70	57.7	230	1	PURQ_METJA	O59042 methanococc
71	57.7	232	1	SCOA_HELPJ	Q9zle3 helicobacte
72	57.7	232	1	SCOA_HELPY	P56006 helicobacte
73	57.7	286	1	CXAC_RAT	P28233 rattus norv
74	57.7	356	1	GBA2_USTMA	P87033 ustilago ma
75	57.7	361	1	RFBB_SALTY	P26391 salmonella

ALIGNMENTS

RESULT 1
CARB_FUSNN
ID CARB_FUSNN STANDARD; PRT; 1058 AA.
AC Q8RG86;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).
GS CARB OR FN0422.

OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;

RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyripides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).

CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP +

```

RC STRAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltyens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Dama R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RT Nature 415:871-880(2002).
RL
CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL031261; CAA20298.1; -.
CC InterPro: IPR002645; STAS.
CC DR InterPro: IPR001902; Sulfate_transp.
CC DR Pfam: PF00916; Sulfate_transp; 1.
CC DR Pfam: PF01740; STAS; 1.
CC DR TIGRfam: TIGR00815; sulp; 1.
CC DR PROSITE: PS01130; SLC26A; 1.
CC DR PROSITE: PS00801; STAS; 1.
CC DR Transport; Transmembrane.
CC KW TRANSMEM 133 153
CC FT TRANSMEM 161 181
CC FT TRANSMEM 186 206
CC FT TRANSMEM 221 241
CC FT TRANSMEM 243 263
CC FT TRANSMEM 292 312
CC FT TRANSMEM 329 349
CC FT TRANSMEM 384 404
CC FT TRANSMEM 424 444
CC FT TRANSMEM 461 481
CC FT TRANSMEM 484 504
CC FT TRANSMEM 518 538
CC FT TRANSMEM 543 563
CC FT DOMAIN 594 747
CC STAS.
CC SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;

Query Match 71.2%; Score 37; DB 1; Length 877;
Best Local Similarity 77.8%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPGXMGSYS 11
    ||| |||||:

```

Db 148 VVPQGMXYA 156

RESULT 3

Y1LK_TYDVA STANDARD; PRT; 102 AA.

AC P31619;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last annotation update)

DE Hypothetical 11.2 kDa protein.

GN V1.

OS Tobacco yellow dwarf virus (strain Australia) (TYDV).

OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.

OX NCBI_TaxID=31599;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92188538; PubMed=1546458;

RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT "The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants.";

RL Virology 187:633-642(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M81103; AAA47947.1; .

DR PIR; A42452; A42452.

DR InterPro; IPR002621; Gemini_mov.

DR Pfam; PF01708; Gemini_mov; 1.

KW Hypothetical protein.

SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 102;

Best Local Similarity 60.0%; Pred. No. 1.4;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 EVVPXGMSYS 11

DB 7 QVVPXGINS 16

RESULT 4

Y1A9_CLOAB STANDARD; PRT; 1498 AA.

AC Q04351;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Hypothetical protein CAC3709.

GN CAC3709.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=1146286;

RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

RA Taturov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RA MEDLINE=93273706; PubMed=8501044;

RA Sauer U., Duerre P.;

RT "Sequence and molecular characterization of a DNA region encoding a small heat shock protein of Clostridium acetobutylicum.";

RL J. Bacteriol. 175:3394-3400(1993).

CC -!- SIMILARITY: BELONGS TO THE FTSK/SPOIIIE FAMILY.

CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts in positions 76 and 106.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE007866; AAK81629.1; .

DR EMBL; X65276; CAA46379.1; ALT_FRAME.

DR InterPro; IPR002543; FtsK_SpoIIIE.

DR Pfam; PF01580; FtsK_SpoIIIE; 2.

KW Hypothetical protein: ATP-binding; Complete proteome.

FT NP_BIND 675 682 ATP (POTENTIAL)

SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 1498;

Best Local Similarity 60.0%; Pred. No. 23;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ERVVPXGMSY 10

DB 1276 EQIPXGMSY 1285

RESULT 5

ZEP1_HUMAN STANDARD; PRT; 2717 AA.

ID ZEP1_HUMAN

AC P15822;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding protein 1) (MBP-1) (Positive regulatory domain II binding factor 1) (PRDII-BF1).

DE (PRDII-BF1).

GN HIVEP1 OR ZNF40.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90169514; PubMed=2106471;

RA Fan C.M., Maniatis T.;

RT "A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence.";

RL Genes Dev. 4:29-42(1990).

RN [2]

RP STRUCTURE BY NMR OF 2113-2142.

RX MEDLINE=91064333; PubMed=2248949;

RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K., Gronenborn A.M.;

RT "High-resolution three-dimensional structure of a single zinc finger from a human enhancer binding protein in solution.";

RL Biochemistry 29:9324-9334(1990).

RN [3]

RP STRUCTURE BY NMR OF 2087-2142.

RX MEDLINE=92232684; PubMed=1567844;

RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E., Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc

finger from the human enhancer binding protein MBP-1.";
 Biochemistry 31:3907-3917(1992).
 -!- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
 NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
 IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
 OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
 MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
 IN T-CELL ACTIVATION.
 -!- SUBCELLULAR LOCATION: Nuclear.
 -!- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
 -!- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
 SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
 ZINC-FINGER IN-BETWEEN.
 -!- SIMILARITY: STRONG, TO HIVP2.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: X51435; CAA35798.1; -
 EPI: A34203; A34203.
 PDB: 3ZNF; 15-JAN-92.
 PDB: 4ZNF; 15-JAN-92.
 PDB: 1BBO; 31-OCT-93.
 TRANSFAC: T00497; -
 Genew: HGNC:4920; HIVEP1.
 MIM: 194540; -
 InterPro: IPR000822; Znf_C2H2.
 Pfam: PF00096; zf-C2H2; 5.
 PRINTS: PR00048; ZINC-FINGER.
 SMART: SM00355; Znf_C2H2; 4.
 PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
 PROSITE: PS00157; ZINC_FINGER_C2H2_2; 4.
 Transcription regulation: Zinc-finger; Metal-binding; DNA-binding;
 Nuclear protein; Repeat; 3D-structure.
 FT DOMAIN 406 456 ZINC FINGERS.
 FT ZN_FING 406 428 ZN_FING
 FT ZN_FING 434 456 C2H2-TYPE.
 FT DOMAIN 803 806 POLY-SER.
 FT ZN_FING 958 981 ZHC-TYPE (POTENTIAL).
 FT DOMAIN 2087 2139 ZINC FINGERS.
 FT ZN_FING 2087 2109 C2H2-TYPE.
 FT ZN_FING 2115 2139 C2H2-TYPE.
 FT STRAND 2088 2088 C2H2-TYPE.
 FT TURN 2090 2092
 FT STRAND 2095 2095
 FT HELIX 2099 2108
 FT TURN 2109 2109
 FT STRAND 2115 2116
 FT STRAND 2123 2124
 FT HELIX 2127 2135
 SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;
 Query Match 67.3%; Score 35; DB 1; Length 2717;
 Best Local Similarity 66.7%; Pred. No. 70;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 VVPXGMSYS 11
 Db 2405 VVPAGLTYS 2413
 RESULT 6
 CY14_NEUCR
 ID CY14_NEUCR STANDARD; PRT; 788 AA.
 AC P23622;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sulfate permease II.
 GN CYS-14.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OC NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91129256; PubMed=1825178;
 RA Ketter J.S., Jarai G., Fu Y.-H., Marzluf G.A.;
 RT "Nucleotide sequence, messenger RNA stability, and DNA recognition
 elements of cys 14," the structural gene for sulfate permease II in
 Neurospora crassa.";
 RL Biochemistry 30:1780-1787(1991).
 RN [2]
 RP PROBABLE REVISIONS.
 RX MEDLINE=94188926; PubMed=8140616;
 RA Sandal N.N., Marcker K.A.;
 RT "Similarities between a soybean nodulin, Neurospora crassa sulphate
 permease II and a putative human tumour suppressor.";
 RL Trends Biochem. Sci. 19:19-19(1994).
 CC -!- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- INDUCTION: Highly expressed, but only in cells subject to sulfur
 limitation, and it is turned on by the positive-acting Cys-3
 sulfur regulatory protein.
 CC -!- MISCELLANEOUS: SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.
 CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: M59167; AAA33615.1; ALT_SEQ.
 PIR: A37956; A37956.
 InterPro: IPR001902; Sulfate_transp.
 Pfam: PF00916; Sulfate_transp; 1.
 TIGRFAMS: TIGR00815; sulp; 1.
 PROSITE: PS01130; SLC26A; 1.
 KW Transport; Transmembrane; Glycoprotein.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 193 213 POTENTIAL.
 FT TRANSMEM 271 291 POTENTIAL.
 FT TRANSMEM 326 346 POTENTIAL.
 FT TRANSMEM 363 383 POTENTIAL.
 FT TRANSMEM 451 471 POTENTIAL.
 FT TRANSMEM 474 494 POTENTIAL.
 FT CARBOHYD 23 23 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 788 AA; 87864 MW; 4FC604B60798CE77 CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 788;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 VVPXGMSYS 11
 Db 90 VVPQGMAYA 98
 RESULT 7
 A10C_HUMAN
 ID A10C_HUMAN STANDARD; PRT; 1499 AA.
 AC O60312; Q96914;
 DT 30-MAY-2000 (Rel. 39, Created)

15-JUN-2002 (Rel. 41, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)
 (Aminophospholipid translocase VC).
 GN ATP10C OR ATPVC OR KIAA0566.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21225279; PubMed=11326269;
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
 RA Oshimura M.;
 RT "A novel maternally expressed gene, ATP10C, encodes a putative
 RT aminophospholipid translocase associated with Angelman syndrome.";
 RN Nat. Genet. 28:19-20(2001).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21313119; PubMed=11353404;
 RA Herzog L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;
 RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
 RT adjacent to UBE3A and exhibits similar imprinted expression.";
 RN Am. J. Hum. Genet. 68:1501-1505(2001).
 RL [3]
 RP SEQUENCE OF 337-1499 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + H₂O = ADP + phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- TISSUE SPECIFICITY: Wide expression, with highest levels in
 CC kidney, followed by lung, brain, prostate, testis, ovary, and
 CC small intestine.
 CC -1- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome
 CC (AS), also known as 'happy puppet syndrome'.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
 CC ATPASES). SUBFAMILY IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB051358; BAB47392.1; -.
 DR EMBL; AY029504; AAK33100.1; -.
 DR EMBL; AY029487; AAK33100.1; JOINED.
 DR EMBL; AY029488; AAK33100.1; JOINED.
 DR EMBL; AY029489; AAK33100.1; JOINED.
 DR EMBL; AY029490; AAK33100.1; JOINED.
 DR EMBL; AY029491; AAK33100.1; JOINED.
 DR EMBL; AY029492; AAK33100.1; JOINED.
 DR EMBL; AY029493; AAK33100.1; JOINED.
 DR EMBL; AY029494; AAK33100.1; JOINED.
 DR EMBL; AY029495; AAK33100.1; JOINED.
 DR EMBL; AY029496; AAK33100.1; JOINED.
 DR EMBL; AY029497; AAK33100.1; JOINED.
 DR EMBL; AY029498; AAK33100.1; JOINED.
 DR EMBL; AY029499; AAK33100.1; JOINED.
 DR EMBL; AY029500; AAK33100.1; JOINED.
 DR EMBL; AY029501; AAK33100.1; JOINED.
 DR EMBL; AY029502; AAK33100.1; JOINED.
 DR EMBL; AY029503; AAK33100.1; JOINED.
 DR EMBL; AB011138; BAA25492.1; -.
 DR EMBL; HGNC:13547; ATP10C.

DR MIM; 605855; -.
 DR MIM; 105830; -.
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR InterPro; IPR001454; Hlgase/hydrase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATAPASE.
 DR PROSITE; PS00154; ATPASE_E1_E2; 1.
 KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KW Multigene family.
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 87 106 POTENTIAL.
 FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 111 128 POTENTIAL.
 FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 310 332 POTENTIAL.
 FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 363 384 POTENTIAL.
 FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1088 1108 POTENTIAL.
 FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1120 1140 POTENTIAL.
 FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1171 1192 POTENTIAL.
 FT DOMAIN 1193 1222 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1200 1222 POTENTIAL.
 FT DOMAIN 1223 1228 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1229 1249 POTENTIAL.
 FT DOMAIN 1250 1267 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1268 1292 POTENTIAL.
 FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).
 FT MOD_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).
 FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).
 FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).
 FT DOMAIN 467 470 POLY-GLU.
 FT CONFLICT 388 388 Q -> R (IN REF. 3).
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 1499;
 Best Local Similarity 72.7%; Pred. No. 62;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 D 469 EEVVPXGMSYS 479
 ID RL20_MYCPU STANDARD; PRT; 116 AA.
 AC Q98QV0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 50S ribosomal protein L20.
 GN RPLT OR MYPU.2610.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAB CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.;
 RT "The complete genome sequence of the murine respiratory pathogen
 RT Mycoplasma pulmonis.";
 RL Nucleic Acids Res. 29:2145-2153(2001).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
 CC NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL
 CC SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS
 CC OF THAT SUBUNIT (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL445563; CAC13434.1; -;
 DR MyuList; MYPU_2610; -;
 DR InterPro; IPR001081; Ribosomal_L20;
 DR Pfam; PF00453; Ribosomal_L20; 1;
 DR PRINTS; PR00062; RIBOSOMALL20;
 DR ProDom; PD002389; Ribosomal_L20; 1;
 DR TIGRFAMs; TIGR01032; rplT_bact; 1;
 DR PROSITE; PS00937; RIBOSOMAL_L20; 1;
 DR Ribosomal protein; rRNA-binding; Complete proteome.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SQ SEQUENCE 116 AA; 13565 MW; C59C748901B1814 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 116;
 Best Local Similarity 77.8%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 DB 68 VRPLGMSYS 76

RESULT 9
 Y990_CAMJE STANDARD; PRT; 253 AA.
 AC P45489; QPNV0; PRT; 253 AA.
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Cj0990C.
 GN Cj0990C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 CC Campylobacter.
 OX NCBI_TaxID=137;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=NCCT 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 RN [2]
 RC SEQUENCE OF 160-253 FROM N.A.
 RX STRAIN=ATCC 43431 / TGH 9011;
 RX MEDLINE=95247673; PubMed=7730270;
 RA Hani E.K., Chan V.L.;
 RA "Expression and characterization of Campylobacter jejuni
 RT benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia
 RT coli.";
 RL J. Bacteriol. 177:2396-2402(1995).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL139076; CAB73246.1; -;

DR EMBL; Z36940; CAA85392.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;

Query Match 63.5%; Score 33; DB 1; Length 253;
 Best Local Similarity 55.6%; Pred. No. 16;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
 DB 185 DIPPSGMSY 193

RESULT 10
 AROA_VIBCH STANDARD; PRT; 426 AA.
 ID AROA_VIBCH
 AC Q9KRE0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN AROA OR VCI732.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC sixth step.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE004251; AAF94882.1; -;
 DR TIGR; VC1732; -;
 DR InterPro; IPR001986; EPSP_synthase.
 DR Pfam; PF00275; EPSP_synthase_1.
 DR ProDom; PD001867; EPSP_synthase; 1.
 DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
 DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
 SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFEC3 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 426;
 Best Local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 DB 223 EFVIPAGQSY 232

DR EMBL; BC010095; AAH10095.1; -
 DR EMBL; AF296124; AAG30413.1; -
 DR EMBL; AL359335; CAB94786.1; -
 DR EMBL; AL359336; CAB94787.1; -
 DR EMBL; AL359336; CAB94787.1; -
 DR EMBL; AL122063; CAB59242.1; -
 DR SWISS-2DPAGE; O9NZM5; HUMAN.
 DR Genew; HGNC:4333; GLUTSR2.
 DR MIM; 605691; -
 KW Nuclear protein; Polymorphism.
 FT VARIANT 389 389 R -> Q.
 FT FTID=VAR_011486.
 FT GGS -> HEG (IN REF. 2; AAH04229).
 FT G -> R (IN REF. 3).
 FT RRKEQLWEKLAQGGELPREVRAQAARLLNPSTRAKPGPOD
 FT TVERP -> SGRSYGRSWPSSRASSPGGAQGPSVAQPFEN
 FT KGNPAPAGHRIAA (IN REF. 3).
 FT SDNPLDRPLVGQDEFFLE -> LNNPDKPVVWPGCLFPG
 FT (IN REF. 3).
 FT A -> S (IN REF. 2; AAH04229).
 FT D -> H (IN REF. 3).
 FT PEGNILDRDFKSFQRNMIEPRERAKFKRYKVKLVEKRAF
 FT REIQ -> VLTVCRCGAPCVMTPTSLPPVPPRGYGRHHGCP
 FT WAGPVGPMPRG (IN REF. 5).
 FT EGNILDRDFKSFQRNMIEPRERAKFKRYKVKLVEKRAF
 FT EIQL -> RGQHSFETGSRAPRGGI (IN REF. 3).
 FT SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;
 Query Match 63.5%; Score 33; DB 1; Length 478;
 Best Local Similarity 60.0%; Pred. No. 3;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 EVVPGXMSYS 11
 |||||
 DB 239 EVAPAGASYN 248
 RESULT 13
 BCN5_CLOPE STANDARD; PRT; 890 AA.
 ID BCN5_CLOPE
 AC P08696;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Bacteriocin BCN5.
 GN BCN.
 OS Clostridium perfringens.
 OG Plasmid pip404.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=88336297; PubMed=2901768;
 RA Garnier T., Cole S.T.;
 RT "Complete nucleotide sequence and genetic organization of the
 RT bacteriocinogenic plasmid, pip404, from Clostridium perfringens.";
 RL Plasmid 19:134-150(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=87057020; PubMed=2877971;
 RA Garnier T., Cole S.T.;
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium
 RT perfringens and molecular genetic analysis of the
 RT bacteriocin-encoding gene.";
 RL J. Bacteriol. 168:1189-1196(1986).
 RN [3]
 RP SEQUENCE OF 1-14 FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=89039249; PubMed=2460717;
 RA Garnier T., Cole S.T.;
 RT "Studies of UV-Inducible promoters from Clostridium perfringens in

RT vivo and in vitro.";
 RL MOL. Microbiol. 2:607-614(1988).
 CC -!- FUNCTION: MAY FUNCTION AS AN IONOPHORE.
 CC -!- INDUCTION: BY UV IRRADIATION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M14481; AAA98248.1; -
 DR EMBL; M32882; AAA98249.1; -
 DR PIR; A30481; A30481.
 DR InterPro: IPR003646; SH3_bac.
 DR SMART; SM00287; SH3b; 3. Plasmid.
 DR Antibiotic; Bacteriocin; Plasmid.
 KW DOMAIN 815 869 HYDROPHOBIC.
 FT SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;
 Query Match 63.5%; Score 33; DB 1; Length 890;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 2 EVVPGXMSYS 10
 |||||
 DB 170 EVVPGGFTY 178
 RESULT 14
 LTBS_HUMAN
 ID LTBS_HUMAN STANDARD; PRT; 1394 AA.
 AC P22064;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Latent transforming growth factor beta binding protein is precursor
 DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-
 DE 1).
 GN LTBP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Fibroblast, and platelet;
 RX MEDLINE=90275601; PubMed=2350783;
 RA Kanzaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;
 RT "TGF-beta 1 binding protein: a component of the large latent complex
 RT of TGF-beta 1 with multiple repeat sequences.";
 RL Cell 61:1051-1061(1990).
 CC -!- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
 CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE
 CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-
 CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a
 CC long form (AC Q14766); are produced by alternative splicing.
 CC -!- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

FT	DISULFID	902	915		BY SIMILARITY.
FT	DISULFID	921	933		BY SIMILARITY.
FT	DISULFID	927	942		BY SIMILARITY.
FT	DISULFID	944	957		BY SIMILARITY.
FT	DISULFID	963	975		BY SIMILARITY.
FT	DISULFID	970	984		BY SIMILARITY.
FT	DISULFID	986	1000		BY SIMILARITY.
FT	DISULFID	1101	1114		BY SIMILARITY.
FT	DISULFID	1109	1123		BY SIMILARITY.
FT	DISULFID	1125	1138		BY SIMILARITY.
FT	DISULFID	1144	1155		BY SIMILARITY.
FT	DISULFID	1150	1164		BY SIMILARITY.
FT	DISULFID	1166	1179		BY SIMILARITY.
FT	DISULFID	1298	1309		BY SIMILARITY.
FT	DISULFID	1304	1318		BY SIMILARITY.
FT	DISULFID	1320	1333		BY SIMILARITY.
FT	DISULFID	1339	1354		BY SIMILARITY.
FT	DISULFID	1349	1363		BY SIMILARITY.
FT	DISULFID	1365	1378		BY SIMILARITY.
FT	MOD_RES	647	647		HYDROXYLATION.
FT	MOD_RES	810	810		HYDROXYLATION.
FT	CARBOHYD	21	21		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	52	52		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	98	98		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	294	294		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	870	870		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	923	923		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1039	1039		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	1394 AA;	152791 MW;	DFFCA81A40B2C7D1 CRC64;	/FTDI-CAR_000184.
QY		1 EEVVPXGMSYS 11			
DB		: : :			
		399 KEICPGMGVT 409			
Query Match		63.5%;	Score 33;	DB 1;	Length 1394;
Best Local Similarity		45.5%;	Pred. No. 93;		
Matches	5;	Conservative	3;	Mismatches	3; Indels 0; Gaps 0;
RESULT 15					
RPOC_VIBCH					
ID	RPOC_VIBCH	STANDARD;	PRT;	1401 AA.	
AC	O9KV29;				
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase				
DE	beta' chain) (RNA polymerase beta' subunit).				
GN	RPOC OR VC0329.				
OS	Vibrio cholerae.				
OC	Bacterii; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_Taxid=666;				
RN	[1]				
SEQUENCE FROM N.A.					
RC	STRAIN=E1 Tor M16961 / Serotype O1;				
RX	MEDLINE=20406833; PubMed=10952301;				
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,				
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
RA	Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,				
RA	McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,				
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,				
RA	Fraser C.M.;				
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio				
RT	cholerae.";				
RL	Nature 406:477-483(2000).				
CC	-!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION				
CC	OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS				
CC	SUBSTRATES (By similarity).				
CC	-!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +				
CC	[RNA] (N).				
CC	-!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE				

FT DISULFID 998 1009 BY SIMILARITY.
 FT DISULFID 1004 1018 BY SIMILARITY.
 FT DISULFID 1020 1033 BY SIMILARITY.
 FT DISULFID 1039 1051 BY SIMILARITY.
 FT DISULFID 1046 1060 BY SIMILARITY.
 FT DISULFID 1062 1074 BY SIMILARITY.
 FT DISULFID 1080 1092 BY SIMILARITY.
 FT DISULFID 1086 1101 BY SIMILARITY.
 FT DISULFID 1103 1116 BY SIMILARITY.
 FT DISULFID 1122 1134 BY SIMILARITY.
 FT DISULFID 1128 1143 BY SIMILARITY.
 FT DISULFID 1145 1158 BY SIMILARITY.
 FT DISULFID 1164 1176 BY SIMILARITY.
 FT DISULFID 1171 1185 BY SIMILARITY.
 FT DISULFID 1201 1215 BY SIMILARITY.
 FT DISULFID 1302 1315 BY SIMILARITY.
 FT DISULFID 1310 1324 BY SIMILARITY.
 FT DISULFID 1326 1339 BY SIMILARITY.
 FT DISULFID 1345 1356 BY SIMILARITY.
 FT DISULFID 1351 1365 BY SIMILARITY.
 FT DISULFID 1367 1380 BY SIMILARITY.
 FT DISULFID 1499 1510 BY SIMILARITY.
 FT DISULFID 1505 1519 BY SIMILARITY.
 FT DISULFID 1521 1534 BY SIMILARITY.
 FT DISULFID 1540 1555 BY SIMILARITY.
 FT DISULFID 1550 1564 BY SIMILARITY.
 FT DISULFID 1566 1579 BY SIMILARITY.
 FT MOD_RES 848 848 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 1011 1011 HYDROXYLATION (BY SIMILARITY).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1124 1124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1240 1240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1595 AA; 173229 MW; 6A091EBA8556D8E5 CRC64;

Query Match 63.58; Score 33; DB 1; Length 1595;
 Best Local Similarity 45.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
 Db 600 KEICPGMGYT 610
 :|: |||:

RESULT 17
 LTBL_RAT STANDARD; PRT; 1712 AA.
 AC Q00918;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE Latent transforming growth factor beta binding protein 1 precursor
 DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-1)
 DE 1) (Transforming growth factor beta-1 masking protein, large subunit).
 GN LTBL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=91062373; PubMed=2247454;
 RA Tsuji T., Okada F., Yamaguchi K., Nakamura T.;
 RT "Molecular cloning of the large subunit of transforming growth factor
 RT type beta masking protein and expression of the mRNA in various rat
 RT tissues.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8835-8839(1990).
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
 CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
 CC A MASKING PROTEIN CONSISTING OF A DISULFIDE-BONDED COMPLEX OF A
 CC DIMER OF THE N-TERMINAL PROPEPTIDE OF THE TGF-BETA1 PRECURSOR AND
 CC A THIRD COMPONENT DENOTED TGF-BETA1-BP (OR MP LARGE SUBUNIT).
 CC TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.

CC -1- SIMILARITY: CONTAINS 18 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M55431; AAA42235.1; --
 DR PIR; A38261; A38261.
 DR HSP; P16109; JFSP.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF-Ca.
 DR InterPro; IPR002212; Fibril-assoc.
 DR Pfam; PF00008; EGF; 16.
 DR Pfam; PF00683; TB; 4.
 DR SMART; SM00179; EGF_Ca; 13.
 DR SMART; SM00001; EGF_Like; 5.
 DR PROSITE; PS00010; ASX_HYDROXYL; 13.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 10.
 DR PROSITE; PS01187; EGF_Ca; 15.
 KW Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 736 POTENTIAL.
 FT CHAIN 737 1577 LATENT TRANSFORMING GROWTH FACTOR BETA
 FT BINDING PROTEIN 1.
 FT SITE 734 736 CLEAVAGE (POTENTIAL).
 FT SITE 1575 1577 CLEAVAGE (POTENTIAL).
 FT PROPEP 1578 1712 CLEAVAGE (POTENTIAL).
 FT DOMAIN 181 213 EGF-LIKE 1.
 FT DOMAIN 391 423 EGF-LIKE 2.
 FT REPEAT 551 604 INTERNAL REPEAT 1.
 FT DOMAIN 618 658 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT REPEAT 671 721 INTERNAL REPEAT 2.
 FT DOMAIN 865 906 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 907 948 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 949 989 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 990 1029 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1030 1070 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1071 1111 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1112 1152 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1153 1193 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1194 1235 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1236 1277 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1278 1320 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT REPEAT 1340 1392 INTERNAL REPEAT 3.
 FT DOMAIN 1415 1457 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1458 1498 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT REPEAT 1517 1568 INTERNAL REPEAT 4.
 FT DOMAIN 1612 1652 EGF-LIKE 17.
 FT DOMAIN 1653 1697 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 185 195 BY SIMILARITY.
 FT DISULFID 199 201 BY SIMILARITY.
 FT DISULFID 203 212 BY SIMILARITY.
 FT DISULFID 395 405 BY SIMILARITY.
 FT DISULFID 399 411 BY SIMILARITY.
 FT DISULFID 413 422 BY SIMILARITY.
 FT DISULFID 622 633 BY SIMILARITY.
 FT DISULFID 628 642 BY SIMILARITY.
 FT DISULFID 644 657 BY SIMILARITY.
 FT DISULFID 869 881 BY SIMILARITY.
 FT DISULFID 876 890 BY SIMILARITY.
 FT DISULFID 892 905 BY SIMILARITY.
 FT DISULFID 911 923 BY SIMILARITY.
 FT DISULFID 918 932 BY SIMILARITY.
 FT DISULFID 934 947 BY SIMILARITY.
 FT DISULFID 953 964 BY SIMILARITY.
 FT DISULFID 959 973 BY SIMILARITY.
 FT DISULFID 976 988 BY SIMILARITY.

FT DISULFID 994 1005 BY SIMILARITY.
 FT DISULFID 1000 1014 BY SIMILARITY.
 FT DISULFID 1017 1028 BY SIMILARITY.
 FT DISULFID 1034 1045 BY SIMILARITY.
 FT DISULFID 1040 1054 BY SIMILARITY.
 FT DISULFID 1056 1069 BY SIMILARITY.
 FT DISULFID 1075 1086 BY SIMILARITY.
 FT DISULFID 1081 1095 BY SIMILARITY.
 FT DISULFID 1097 1110 BY SIMILARITY.
 FT DISULFID 1116 1127 BY SIMILARITY.
 FT DISULFID 1122 1136 BY SIMILARITY.
 FT DISULFID 1138 1151 BY SIMILARITY.
 FT DISULFID 1157 1169 BY SIMILARITY.
 FT DISULFID 1164 1178 BY SIMILARITY.
 FT DISULFID 1180 1192 BY SIMILARITY.
 FT DISULFID 1198 1210 BY SIMILARITY.
 FT DISULFID 1204 1219 BY SIMILARITY.
 FT DISULFID 1221 1234 BY SIMILARITY.
 FT DISULFID 1240 1252 BY SIMILARITY.
 FT DISULFID 1246 1261 BY SIMILARITY.
 FT DISULFID 1263 1276 BY SIMILARITY.
 FT DISULFID 1282 1294 BY SIMILARITY.
 FT DISULFID 1289 1303 BY SIMILARITY.
 FT DISULFID 1305 1319 BY SIMILARITY.
 FT DISULFID 1419 1432 BY SIMILARITY.
 FT DISULFID 1427 1441 BY SIMILARITY.
 FT DISULFID 1443 1456 BY SIMILARITY.
 FT DISULFID 1462 1473 BY SIMILARITY.
 FT DISULFID 1468 1482 BY SIMILARITY.
 FT DISULFID 1484 1497 BY SIMILARITY.
 FT DISULFID 1616 1627 BY SIMILARITY.
 FT DISULFID 1622 1636 BY SIMILARITY.
 FT DISULFID 1638 1651 BY SIMILARITY.
 FT DISULFID 1657 1672 BY SIMILARITY.
 FT DISULFID 1667 1681 BY SIMILARITY.
 FT DISULFID 1683 1696 BY SIMILARITY.
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1357 1357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1712 AA; 186598 MW; 650BCBAA691FD134 CRC64;
 Query Match 63.5%; Score 33; DB 1; Length 1712;
 Best Local Similarity 45.5%; Pred. No. 1.le+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSYS 11
 Db 718 KEICPGMGYT 728
 RESULT 18
 CGD2_RAT
 ID CGD2_RAT STANDARD; PRT; 288 AA.
 AC Q04827;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GL/S-specific cyclin D2 (Vin-1 proto-oncogene).
 GN CCND2 OR VIN-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93275661; PubMed=8502486;
 RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,
 RA Francke U., Jolicoeur P.;
 RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is

RT the cyclin D2.";
 RL Oncogene 8:1661-1666(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95011623; PubMed=7926809;
 RA Hosokawa Y., Onga T., Nakashima K.;
 RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the
 RT G1/S transition by prolactin in rat Nb2 cells.";
 RL Gene 147:249-252(1994).
 CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
 CC (START) TRANSITION
 CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
 CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L09752; AAA41010.1; -;
 DR EMBL; D16308; BAA03815.1; -;
 DR InterPro; IPR004366; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR Pfam; PF00134; cyclin; 1.
 DR Pfam; PF02984; cyclin_C; 1.
 DR SMART; SM00385; CYCLIN; 1.
 DR PROSITE; PS00292; CYCLINS; 1.
 KW Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.
 FT CONFLICT 68 68 E -> G (IN REF. 2).
 FT CONFLICT 104 104 C -> V (IN REF. 2).
 FT CONFLICT 232 232 T -> A (IN REF. 2).
 SQ SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;
 Query Match 61.5%; Score 32; DB 1; Length 288;
 Best Local Similarity 60.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGMSY 10
 Db 73 EEVFPPLANNY 82
 RESULT 19
 CGD2_HUMAN
 ID CGD2_HUMAN STANDARD; PRT; 289 AA.
 AC P30279; Q13955;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GL/S-specific cyclin D2.
 GN CCND2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347851; PubMed=1386336;
 RA Xiong Y., Menninger J., Beach D., Ward D.C.;
 RT "Molecular cloning and chromosomal mapping of CCND genes encoding
 RT human D-type cyclins.";
 RL Genomics 13:575-584(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93205384; PubMed=8455931;
 RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
 RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
 RT cell lines.";

```

RL Oncogene 8:1049-1054(1993).
RN [3]
RP SEQUENCE FROM N.A.
RA Miyajima N.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-240 FROM N.A.
RC TISSUE=Placenta;
RA Inaba T., Matsushime H., Valentine M., Roussel M.F., Sherr C.J.,
  Look A.T.;
  genomic organization, chromosomal localization, and independent
  expression of human cyclin D genes."
  Genomics 13:565-574(1992).
RL Genomics 13:565-574(1992).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
  (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
  A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
  IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90813; AAA51926.1; -
DR EMBL; D68452; CAA48493.1; -
DR EMBL; D13639; BAA02802.1; -
DR EMBL; BC010958; RAH10958.1; -
DR EMBL; M88083; AAA51928.1; -
DR EMBL; M88080; AAA51928.1; JOINED.
DR EMBL; M88081; AAA51928.1; JOINED.
DR EMBL; M88082; AAA51928.1; JOINED.
DR PIR; A42822; A42822.
DR PIR; S26580; S26580.
DR Genew; HGNC:1583; CCND2.
DR MIM; 123833; -
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
FT CONFLICT 166 167 KL -> NV (IN REF. 5).
FT CONFLICT 224 224 T -> H (IN REF. 5).
SQ SEQUENCE 289 AA; 33067 MW; E4E5FEF476D76D90 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFPPLANNY 83

RESULT 20
CGD2_MOUSE
ID CGD2_MOUSE STANDARD; PRT; 289 AA.
AC P30280;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE G1/S-specific cyclin D2.
GN CCND2 OR CYL-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196134; PubMed=1372445;
RA Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,
  Marks P.A.;
  "Cloning of a D-type cyclin from murine erythroleukemia cells."
  Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91235305; PubMed=1827757;
RA Matsushime H., Roussel M.F., Ashmun R.A., Sherr C.J.;
  "Coloncy-stimulating factor 1 regulates novel cyclins during the G1
  phase of the cell cycle."
  Cell 65:701-713(1991).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
  (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
  A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
  IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M83749; AAA37519.1; -
DR EMBL; M86182; AAA37503.1; -
DR PIR; B40035; B40035.
DR PIR; A41984; A41984.
DR MGD; MGI:88314; Ccnd2.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
KW SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;

Query Match 61.5%; Score 32; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 73 EEVFPPLANNY 82

RESULT 21
CGD1_BRARE
ID CGD1_BRARE STANDARD; PRT; 291 AA.
AC Q90459;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D1.
GN CGD1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96138542; PubMed=8547308;
RA Yarden A., Salomon D., Geiger B.;
RT "zebrafish cyclin D1 is differentially expressed during early
embryogenesis.";
RL Biochim. Biophys. Acta 1264:257-260(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
(START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X87581; CAA60885.1; -
DR ZFIN; ZDB-GENE-980526-176; cycl1.
DR InterPro; IPR004366; Cyclin.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division.
KW SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;
SQ
Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
DB 75 EEVFPPLANNY 84
RESULT 22
CGDI_XENLA
ID CGDI_XENLA STANDARD; PRT; 291 AA.
AC P50755;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D1.
GN CCND1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cockerill M.J., Hunt T.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
(START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

CC EMBL; X89475; CAA61664.1; -
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
KW SEQUENCE 291 AA; 32953 MW; A4747C5BD1679087 CRC64;
SQ
Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
DB 73 EEVFPPLANNY 82
RESULT 23
CGD2_CHICK
ID CGD2_CHICK STANDARD; PRT; 291 AA.
AC P49706;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D2.
GN CCND2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144302; PubMed=8566807;
RA Li H., Grenet J., Kidd V.J.;
RT "Structure and gene expression of avian cyclin D2.";
RL Gene 167:341-342(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
(START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U28980; AAA96955.1; -
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
KW SEQUENCE 291 AA; 33163 MW; 43A7E646AEAF3109 CRC64;
SQ
Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
DB 74 EEVFPPLANNY 83


```
RESULT 24
CGD2_XENLA          STANDARD;          PRT;    291 AA.
AC  P33782;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  GI/S-specific cyclin D2.
GN  CCND2.
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC  Xenopodinae; Xenopus.
OX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Cockerill M.J., Hunt T.;
RL  Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
[2]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=97380591; PubMed=9237366;
RA  Taieb F., Jessus C.; cloning and expression in oocytes and during
RT  early development.;
RL  Biol. Cell 88:99-111(1996).
CC  -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC  (START) TRANSITION.
CC  -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC  A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC  IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC  -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL; X89476; CAA61665.1; -
EMBL; X83503; CAA58493.1; -
DR  InterPro; IPR004366; Cyclin.
DR  InterPro; IPR004367; Cyclin_Cterm.
DR  Pfam; PF00134; cyclin; 1.
DR  Pfam; PF02984; cyclin_C; 1.
DR  SMART; SM00385; CYCLIN; 1.
DR  PROSITE; PS00292; CYCLINS; 1.
KW  Cyclin; Cell cycle; Cell division; Multigene family.
SQ  SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;

Query Match      61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEVVPXGMSY 10
DB  ||||| |
    74 EEVFPAMNYY 83

RESULT 25
CGD1_CHICK
ID  CGD1_CHICK          STANDARD;          PRT;    292 AA.
AC  P35169;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  GI/S-specific cyclin D1.
GN  CCND1.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Gallus.
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
Li H., Lahti J.M., Kidd V.J.;
Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
(START) TRANSITION. INTERACTS WITH THE CDK4 AND CDK6 PROTEIN
KINASES (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL; U40844; AAA83271.1; -
DR  InterPro; IPR004366; Cyclin.
DR  InterPro; IPR004367; Cyclin_Cterm.
DR  Pfam; PF00134; cyclin; 1.
DR  Pfam; PF02984; cyclin_C; 1.
DR  SMART; SM00385; CYCLIN; 1.
DR  PROSITE; PS00292; CYCLINS; 1.
KW  Cyclin; Cell cycle; Cell division; Multigene family.
SQ  SEQUENCE 292 AA; 33263 MW; 7B543029DB45A67D CRC64;

Query Match      61.5%; Score 32; DB 1; Length 292;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEVVPXGMSY 10
DB  ||||| |
    75 EEVFPAMNYY 84

Search completed: June 10, 2003, 13:40:25
Job time : 4.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search., using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds
(without alignments)
87.898 Million cell updates/sec

Title: US-09-909-164-47
Perfect score: 52
Sequence: 1 BEVVPXGMSYS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organella:*
9: sp.phage:*
10: sp.plant:*
11: sp.prodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	387	16 Q98FX1	Q98fx1 rhizobium 1
2	38	73.1	1063	16 Q8RG86	Q8rg86 fusobacteri
3	38	73.1	3472	1 Q74056	O74056 cenarchaeum
4	37	71.2	840	3 Q9UR88	Q9ury8 schizosacch
5	36	69.2	471	11 Q8RI26	Q8ri26 mus musculu
6	36	69.2	484	11 Q8VD18	Q8vd18 mus musculu
7	35	67.3	225	10 Q40129	Q40129 lycopersico
8	35	67.3	425	5 Q9XVK4	Q9xvk4 caenorhabdi
9	35	67.3	556	4 Q43733	Q43733 homo sapien
10	35	67.3	583	5 Q9BHA5	Q9bha5 plasmodium
11	35	67.3	583	5 Q9BH83	Q9bh83 plasmodium
12	35	67.3	670	11 Q01487	Q01487 rattus norv
13	35	67.3	749	16 Q9PDM6	Q9pdm6 xylella fas
14	35	67.3	1902	4 Q14122	Q14122 homo sapien
15	34	65.4	156	3 Q12479	Q12479 saccharomyc
16	34	65.4	219	17 Q971S2	Q971s2 sulfolobus

17	34	65.4	252	17	Q28342	O28342 archaeoglob
18	34	65.4	290	4	Q96MU1	Q96mu1 homo sapien
19	34	65.4	387	16	Q92MD6	Q92md6 rhizobium m
20	34	65.4	541	16	Q98BP5	Q98bp5 rhizobium 1
21	34	65.4	544	16	Q9PQD2	Q9pqd2 ureaplasma
22	34	65.4	842	3	Q9URR4	Q9urr4 penicillium
23	34	65.4	1049	16	Q8XT05	Q8xt05 ralstonia s
24	34	65.4	1499	4	Q96914	Q96914 homo sapien
25	33	63.5	143	17	Q8TX62	Q8tx62 methanopyru
26	33	63.5	162	11	Q9CX04	Q9cxq4 mus musculu
27	33	63.5	165	17	Q28330	Q28330 archaeoglob
28	33	63.5	193	2	Q8VUA8	Q8vuua8 lactococcu
29	33	63.5	209	16	Q8RE56	Q8re56 fusobacteri
30	33	63.5	284	16	P74187	P74187 synechocyst
31	33	63.5	298	10	Q9M3C0	Q9m3c0 arabidopsis
32	33	63.5	326	12	Q9Q9Q9	Q9q9q9 soil-borne
33	33	63.5	326	12	Q9Q9Q5	Q9q9q5 soil-borne
34	33	63.5	326	12	Q9QCE7	Q9qce7 soil-borne
35	33	63.5	326	12	Q9DJG4	Q9djg4 soil-borne
36	33	63.5	326	12	Q91DN1	Q91dn1 soil-borne
37	33	63.5	326	12	Q9Q9Q7	Q9q9q7 soil-borne
38	33	63.5	327	12	Q9G360	Q9g360 soil-borne
39	33	63.5	332	10	Q9FNL4	Q9fnl4 arabidopsis
40	33	63.5	368	16	Q9X003	Q9x0u3 thermotoga
41	33	63.5	393	5	Q9V914	Q9v914 drosophila
42	33	63.5	479	4	Q96CS0	Q96cs0 homo sapien
43	33	63.5	548	11	Q9D2X9	Q9d2x9 mus musculu
44	33	63.5	584	16	Q8R8K6	Q8r8k6 thermoanaer
45	33	63.5	648	4	Q96MB2	Q96mb2 homo sapien
46	33	63.5	653	16	Q9KVE3	Q9kve3 vibrio chol
47	33	63.5	676	5	Q9VA55	Q9va55 drosophila
48	33	63.5	676	5	Q8T8Z7	Q8t8z7 drosophila
49	33	63.5	678	12	Q9ELX6	Q9elx6 cercopithe
50	33	63.5	746	3	Q9URR3	Q9urr3 penicillium
51	33	63.5	791	4	Q9H2K5	Q9h2k5 homo sapien
52	33	63.5	793	4	Q9H2K6	Q9h2k6 homo sapien
53	33	63.5	844	11	Q922D4	Q922d4 mus musculu
54	33	63.5	1028	16	Q8VJ11	Q8vj11 brucella me
55	33	63.5	1152	16	Q9CC95	Q9cc95 mycobacteri
56	33	63.5	1305	5	Q9V7C7	Q9v7c7 drosophila
57	33	63.5	1394	4	Q8TD95	Q8td95 homo sapien
58	33	63.5	1442	17	Q96YH5	Q96yh5 sulfolobus
59	33	63.5	1548	10	Q95531	Q95531 arabidopsis
60	33	63.5	1713	11	Q88349	Q88349 mus musculu
61	32	61.5	84	16	Q97DE7	Q97de7 clostridium
62	32	61.5	103	11	Q9D0H9	Q9d0h9 mus musculu
63	32	61.5	108	1	Q9UX33	Q9ux33 sulfolobus
64	32	61.5	153	13	P79919	P79919 xenopus lae
65	32	61.5	156	11	Q9D8L9	Q9d8l9 mus musculu
66	32	61.5	174	10	Q9M3T4	Q9m3t4 betula verr
67	32	61.5	190	13	Q57481	Q57481 stizostedio
68	32	61.5	191	11	Q9NBA4	Q9nba4 rattus norv
69	32	61.5	200	17	Q97CD0	Q97cd0 thermoplasm
70	32	61.5	207	2	Q47284	Q47284 escherichia
71	32	61.5	234	2	O32330	O32330 clostridium
72	32	61.5	236	10	Q9SXF1	Q9sxf1 arabidopsis
73	32	61.5	240	11	Q9DB09	Q9db09 mus musculu
74	32	61.5	243	12	Q91EW1	Q91ew1 cydia pomon
75	32	61.5	279	17	Q9F8Z4	Q9y8z4 aeropyrum p

ALIGNMENTS

RESULT 1
Q98FX1 PRELIMINARY; PRT; 387 AA.
ID Q98FX1
AC Q98FX1
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hippurate hydrolase.
GN MLR3583.

```

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003002; BAB50445.1; -.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
DR Hydrolase; Complete proteome.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 387;
Best Local Similarity 60.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVPXGMSY 10
Db 367 DEAIPIHGMYS 376

RESULT 2
ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
AC Q8RG86;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kypides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010534; AAL94625.1; -.
RW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMSY 11
Db 195 EIVPGLNYS 204

RESULT 3
ID Q74056 PRELIMINARY; PRT; 3472 AA.
AC Q74056;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)

```

```

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 367.1 kDa protein.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., Delong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon 'Cenarchaeum symbiosum.'";
RL J. Bacteriol. 180:5003-5009(1998).
CC 1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AF083072; AAC62699.1; -.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;

Query Match 73.1%; Score 38; DB 1; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVVPXGMSY 11
Db 2294 EDVIPRGISFS 2304

RESULT 4
ID Q9URY8 PRELIMINARY; PRT; 840 AA.
AC Q9URY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable sulfate permease.
GN SPAC869.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI32779; CAB60015.1; -.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulP; 1.
SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSY 11
Db 135 VVPOGMSYA 143

RESULT 5
ID Q8R126 PRELIMINARY; PRT; 471 AA.

```

AC Q8R126;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 54.5 kDa protein (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC025810; AAH25810.1;
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 471;
 Best Local Similarity 60.0%; Pred. No. 32;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
 :||| | :||
 Db 226 EVIPAGASYN 235

RESULT 6
 Q8VD18
 ID Q8VD18 PRELIMINARY; PRT; 484 AA.
 AC Q8VD18;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to glioma tumor suppressor candidate region gene 2.
 GN AW536441.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SALIVARY GLAND;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC017637; AAH17637.1;
 DR MGD; MGI:2138595; AW536441.
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 484;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
 :||| | :||
 Db 239 EVIPAGASYN 248

RESULT 7
 Q40129
 ID Q40129 PRELIMINARY; PRT; 225 AA.
 AC Q40129;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 25.2 kDa protein precursor.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=VF36; TISSUE=PISTIL;
 RX MEDLINE=95375233; PubMed=7647301;
 RA Milligan S.B.; Gasser C.S.;
 RT "Nature and regulation of pistil-expressed genes in tomato.";
 RL Plant Mol. Biol. 28:691-711(1995).
 DR EMBL; U20592; AAA80497.1;
 DR InterPro; IPR002160; Kunitz_legume.
 DR Pfam; PF00197; Kunitz_legume; 1.
 DR ProDom; PD000891; Kunitz_legume; 1.
 DR SMART; SM00452; ST1; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; UNKNOWN_1.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 225 UNKNOWN.
 SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;

Query Match 67.3%; Score 35; DB 10; Length 225;
 Best Local Similarity 54.5%; Pred. No. 24;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSYS 11
 :||| | :||
 Db 32 DEVVPGKTYA 42

RESULT 8
 Q9XVK4
 ID Q9XVK4 PRELIMINARY; PRT; 425 AA.
 AC Q9XVK4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 GN R10D12.10 protein.
 GN R10D12.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Percy C.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81109; CAB03241.1;
 DR InterPro; IPR000719; Euk_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; transferase.
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 425;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10
 :||| | :||
 Db 335 EQIVPGGLQY 344

RESULT 9
 O43733
 ID O43733 PRELIMINARY; PRT; 556 AA.
 AC O43733;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-1998 (Tremblrel. 06, Last annotation update)
 DE DNA binding protein (Fragment).
 GN DJ451B15.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tubby B.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z98050; CAB10847.1; -;
 FT NON_TER 1
 SQ SEQUENCE 556 AA; 59059 MW; ECB00E4033FB2528 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 556;
 Best Local Similarity 66.7%; Pred. No. 64;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 ||| |::||
 Db 244 VVPAGLTYS 252

RESULT 10
 Q9BHA5 PRELIMINARY; PRT; 583 AA.
 AC Q9BHA5;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Choline transporter.
 GN SC1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007372; AAK14816.1; -;
 DR EMBL; AY007375; AAG17947.1; -;
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;
 Best Local Similarity 55.6%; Pred. No. 67;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 ::| |::||
 Db 227 IIPVGLSYS 235

RESULT 11
 Q9BH83 PRELIMINARY; PRT; 583 AA.
 AC Q9BH83;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Choline transporter.
 GN SC1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY007374; AAK14818.1; -;
 DR EMBL; AY007373; AAK14817.1; -;
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;
 Best Local Similarity 55.6%; Pred. No. 67;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 ::| |::||
 Db 227 IIPVGLSYS 235

RESULT 12
 Q01487 PRELIMINARY; PRT; 670 AA.
 ID Q01487;
 AC Q01487;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).
 DE Rattus norvegicus (Rat).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=THYROID;
 RX MEDLINE-91187610; PubMed-1901405;
 RA Mitchelmore C., Traboni C., Cortese R.;
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer.";
 RT Nucleic Acids Res. 19:141-147(1991).
 CC -!- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER; IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.
 CC -!- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.
 DR EMBL; X54250; CAA38151.1; -;
 DR HSP; P15822; 1BBO.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;
 KW Metal-binding; Multigene family.
 FT NON_TER 1
 FT DOMAIN 54 104 ZINC-FINGERS.
 FT DOMAIN 140 160 ACIDIC.
 FT ZN_FING 54 74 C(2)H(2) CLASS.
 FT ZN_FING 82 104 C(2)H(2) CLASS.
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;
 Best Local Similarity 66.7%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
 ||| |::||
 Db 376 VVPAGLTYS 384

RESULT 13

Q9PDM6
ID Q9PDM6 PRELIMINARY; PRT; 749 AA.
AC Q9PDM6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Topoisomerase IV subunit.
GN XF1353.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Canargo A.A., Canargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facchini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshakho M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AF003967; AAF84162.1;
DR HSSP; P09097; 1AB4.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
DR SMART; SM00434; TOP4c; 1.
DR TIGRFAMs; TIGR01062; parC_Gneg; 1.
KW Complete proteome.
SQ SEQUENCE 749 AA; 83344 MW; 4DCD10F480EE0257 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 749;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EYVPGMSY 10
Db 526 EYDPSGMSY 534

RESULT 14
Q14122
ID Q14122 PRELIMINARY; PRT; 1902 AA.
AC Q14122;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-2002 (TREMBLrel. 20, Last annotation update)
DE DNA-binding protein (Mbp-1) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205917; PubMed=2108316;
RA Baldwin A.S., Leclair K.P., Singh H., Sharp P.A.;
RT "A large protein containing zinc finger domains binds to related
RT sequence elements in the enhancers of the class I major
RT histocompatibility complex and kappa immunoglobulin genes.";
RL Mol. Cell. Biol. 10:1406-1414(1990).
DR EMBL; M32019; AAA17534.1;
DR HSSP; P15822; 1BBO.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW DNA-binding; Metal-binding; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E6ED20 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 1902;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
Db 1590 VVPAGLTYS 1598

RESULT 15
Q12479
ID Q12479 PRELIMINARY; PRT; 156 AA.
AC Q12479;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA De Haan M., Grivell L.A., Maarse A.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RA De Haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RT Sherman F.;
RT "CYC2 encodes a factor involved in mitochondrial import of yeast
RT cytochrome c.";
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RT Biosci. Biotechnol. Biochem. 58:391-395(1994).

```

DR EMBL; Z74920; CAA99201.1; -
DR EMBL; X87331; CAA60762.1; -
DR SGD; S0005539; YOR013W;
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 3; Length 156;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
||:| |
DB 50 EYVPLGMDY 58

RESULT 16
Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative ribose 5-phosphate isomerase.
GN ST1302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000985; BAB6348.1; -
DR InterPro; IPR004788; RPIA.
DR ProDom; PD005813; RPIA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 17; Length 219;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
|||| |
DB 131 EVVPVGVA 139

RESULT 17
O28342 PRELIMINARY; PRT; 252 AA.
AC O28342;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cell division inhibitor (MIND-2).
GN AF1937.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE000970; AAB89318.1; -
DR TIGR; AF1937; -
DR InterPro; IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
KW Hypothetical protein; Cell division; Complete proteome.
SQ SEQUENCE 252 AA; 27130 MW; A401DC1F93E8C538 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 17; Length 252;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9
||:| |
DB 81 EVIPAGMS 88

RESULT 18
Q96MU1 PRELIMINARY; PRT; 290 AA.
ID Q96MU1;
AC Q96MU1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CDNA FL31891 fis, clone NT2RP7003304, weakly similar to YceA protein
DE homolog ybfQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056453; BAB71188.1; -
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
SQ SEQUENCE 290 AA; 32972 MW; 41FB7FB0217C1421 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 4; Length 290;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9
||:| |
DB 35 EEIVPMGIS 43

RESULT 19
Q92MD6 PRELIMINARY; PRT; 387 AA.
ID Q92MD6;
AC Q92MD6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```



```

01-DEC-2001 (Tremblrel. 19, Last sequence update)
01-MAR-2002 (Tremblrel. 20, Last annotation update)
Putative HIPPURATE hydrolase protein (EC 3.5.1.32).
GN HPO1 OR R02690 OR SMC00682.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396507; PubMed=11491430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampsperger F.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47269.1; -.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41074 MW; 342763088907A6E3 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 387;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
: | : | : |
Db 367 DEAIHPGISY 376

RESULT 20
Q98BP5 PRELIMINARY; PRT; 541 AA.
AC Q98BP5;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Probable DNA ligase.
GN MLL5481.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51927.1; -.
DR InterPro; IPR000977; DNA_ligase.
DR Pfam; PF01068; DNA_ligase; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 541 AA; 60645 MW; 2E9EF705453F28F8 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 541;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

```

```

Db 445 EELVPVGKAY 454
: | : | : | : |
: | : | : | : |

RESULT 21
Q9PQD2 PRELIMINARY; PRT; 544 AA.
AC Q9PQD2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Putative ABC substrate-binding protein-iron.
GN ABCSEP-5 OR UU359.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROVAR 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,
RA Casseil G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RL Nature 407:757-762(2000).
DR EMBL; AE002133; AAF30768.1; -.
KW Complete proteome.
SQ SEQUENCE 544 AA; 61291 MW; CF8756202A389C00 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 544;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
: | : | : | : |
: | : | : | : |

Db 135 EEVVPHYLSY 144

RESULT 22
Q9URR4 PRELIMINARY; PRT; 842 AA.
AC Q9URR4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sulfate permease SUTB.
GN SUTB.
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Q176;
RX MEDLINE=20042342; PubMed=10572125;
RA Van de Kamp M., Pizzinini E., Vos A., Van der Lende T.R.,
RA Schuurs T.A., Newbert R.W., Turner G., Konings W.N., Driessen A.J.M.;
RT "Sulfate Transport in Penicillium chrysogenum: Cloning and
RT Characterization of the sutA and sutB Genes."
RL J. Bacteriol. 181:7228-7234(1999).
DR EMBL; AF163974; AAF14539.1; -.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF01740; STAS; 1.
DR TIGRFAMs; TIGR00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SULFATE_TRANS; UNKNOWN_1.
SQ SEQUENCE 842 AA; 91865 MW; 839A55486E73D15 CRC64;

Query Match 65.4%; Score 34; DB 3; Length 842;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      3 VVPXGMSYS 11
Db      111 VVPOGMAYA 119

RESULT 23
Q8XT05  ID Q8XT05      PRELIMINARY;      PRT; 1049 AA.
AC      Q8XT05;
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Probable multidrug efflux system transmembrane protein.
GN      MEXD OR RSP0312 OR RS05457.
OS      Ralstonia solanacearum (Pseudomonas solanacearum).
OC      Plasmid megaplasmid.
OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC      Ralstonia.
OX      NCBI_TaxID=305;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21681879; PubMed=11823852;
RA      Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA      Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA      Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA      Gaspin C., Lavie M., Molisan A., Robert C., Saurin W., Schlex T.,
RA      Siquier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA      Welssenbach J., Boucher C.A.;
RA      "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL      Nature 415:497-502(2002).
DR      EMBL; AL646078; CAD17463.1; -.
DR      InterPro; IPR001036; Acrlflin_res.
DR      InterPro; IPR004764; HAE1.
DR      InterPro; IPR000731; HMGCR/patch_5TM.
DR      Pfam; PF00873; ACR_tran; 1.
DR      PRINTS; PR00702; ACRIFLAVINRP.
DR      TIGRFAMs; TIGR00915; 2A0602; 1.
DR      PROSITE; PS01556; SSD; 1.
KW      Plasmid; Complete proteome.
SQ      SEQUENCE 1049 AA; 111769 MW; CB59674B670089CE CRC64;

Query Match      65.4%; Score 34; DB 16; Length 1049;
Best Local Similarity 75.0%; Pred. NO. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 VVPXGMSYS 11
Db      317 MPAGMSYS 324

RESULT 24
Q96914  ID Q96914      PRELIMINARY;      PRT; 1499 AA.
AC      Q96914;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Putative aminophospholipid translocase (Aminophospholipid-transporting
DE      ATPase).
GN      ATP10C.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21225279; PubMed=11326269;
RA      Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
RA      Oshimura M.;
RT      "A novel maternally expressed gene, ATP10C, encodes a putative
RT      aminophospholipid translocase associated with Angelman syndrome.";

```

```

RL      Nat. Genet. 28:19-20(2001).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21313119; PubMed=11353404;
RA      Herzog L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;
RT      "The human aminophospholipid-transporting ATPase gene ATP10C maps
RT      adjacent to UBE3A and exhibits similar imprinted expression.";
RL      Am. J. Hum. Genet. 68:1501-1505(2001).
DR      EMBL; AB051358; BAB47392.1; -.
DR      EMBL; AY029504; AAK33100.1; -.
DR      EMBL; AY029487; AAK33100.1; JOINED.
DR      EMBL; AY029488; AAK33100.1; JOINED.
DR      EMBL; AY029489; AAK33100.1; JOINED.
DR      EMBL; AY029490; AAK33100.1; JOINED.
DR      EMBL; AY029491; AAK33100.1; JOINED.
DR      EMBL; AY029492; AAK33100.1; JOINED.
DR      EMBL; AY029493; AAK33100.1; JOINED.
DR      EMBL; AY029494; AAK33100.1; JOINED.
DR      EMBL; AY029495; AAK33100.1; JOINED.
DR      EMBL; AY029496; AAK33100.1; JOINED.
DR      EMBL; AY029497; AAK33100.1; JOINED.
DR      EMBL; AY029498; AAK33100.1; JOINED.
DR      EMBL; AY029499; AAK33100.1; JOINED.
DR      EMBL; AY029500; AAK33100.1; JOINED.
DR      EMBL; AY029501; AAK33100.1; JOINED.
DR      EMBL; AY029502; AAK33100.1; JOINED.
DR      InterPro; IPR001757; ATPase_E1-E2.
DR      InterPro; IPR001064; Crystallin.
DR      InterPro; IPR001454; Hlgnase/hydrolase.
DR      Pfam; PF00702; Hydrolase; 1.
DR      PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
DR      PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
SQ      SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match      65.4%; Score 34; DB 4; Length 1499;
Best Local Similarity 72.7%; Pred. NO. 3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSYS 11
Db      469 EEVVPXGMSYS 479

RESULT 25
Q8TX62  ID Q8TX62      PRELIMINARY;      PRT; 143 AA.
AC      Q8TX62;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Uncharacterized conserved protein.
GN      MK0814.
OS      Methanopyrus kandleri.
OC      Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC      Methanopyrus.
OX      NCBI_TaxID=2320;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=AV19 / DSM 6324 / JCM 9639;
RX      MEDLINE=21927647; PubMed=11930014;
RA      Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA      Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA      Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA      Malykh A.G., Koonin E.V., Kozlyavkin S.A.;
RT      "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT      and monophyly of archaeal methanogens.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR      EMBL; AE010372; AAK02027.1; -.
KW      Complete proteome.
SQ      SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBEDD0B CRC64;

Query Match      63.5%; Score 33; DB 17; Length 143;

```

Wed Jun 11 15:43:57 2003

us-09-909-164-47.rspt

Page 9

Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 ||:| | | |
Db 75 EELVPQAGY 84

Search completed: June 10, 2003, 13:46:39
Job time : 25.7857 secs

